



## Les matrices : formes de représentations et pratiques opératoires (1850-1930).

Frederic Brechenmacher

### ► To cite this version:

Frederic Brechenmacher. Les matrices : formes de représentations et pratiques opératoires (1850-1930).. Site expert des Ecoles Normales Supérieures., 2006, <http://www.dma.ens.fr/culturemath/hal-00142789>

**HAL Id: hal-00142789**

**<https://hal.science/hal-00142789>**

Submitted on 22 Apr 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Les matrices : formes de représentation et pratiques opératoires (1850-1930).

**Frédéric Brechenmacher** - Centre Alexandre Koyré

---

Article déposé le 20 décembre 2006. Toute reproduction pour publication ou à des fins commerciales, de la totalité ou d'une partie de l'article, devra impérativement faire l'objet d'un accord préalable avec l'éditeur (ENS Ulm). Toute reproduction à des fins privées, ou strictement pédagogiques dans le cadre limité d'une formation, de la totalité ou d'une partie de l'article, est autorisée sous réserve de la mention explicite des références éditoriales de l'article.

---

## SOMMAIRE

Introduction.....	2
I. La théorie des matrices canoniques des années trente du XXe siècle.....	4
1. La nouveauté de la théorie des matrices des années 1920-1930.....	4
2. Les problématiques liées à l'histoire des procédés matriciels.....	7
3. La méthodologie des réseaux de textes.....	8
II. Une première origine : les travaux de Sylvester et Cayley des années 1850.....	10
1. Origine des matrices comme mères des mineurs : les travaux de Sylvester [1850-1851] sur les intersections de deux coniques.....	10
2. La théorie des matrices de Cayley.....	16
III. Evolutions de la notion de matrice entre 1850 et 1890.....	23
1. Héritages de la théorie des matrices de Cayley dans les travaux de Sylvester de la période 1882-1885.....	23
2. La formation des "espèces de matrices" d'Eduard Weyr.....	29
3. La rencontre de la théorie des formes bilinéaires et de la théorie des matrices.....	32
Conclusion.....	37
ENCARTS.....	43
Encart 1 : an introduction to the theory of canonical matrices. <i>Turnbul et Aitken</i> , 1932.....	43
Encart 2 : le théorème de décomposition matricielle dans le traité de Mac Duffee de 1943.....	44
Encart 3 : extraits de la démonstration du théorème de Jordan dans le traité de Mac Duffee de 1943.....	45
Encart 4 : un exemple de représentation de réseaux de recherches (1880-1907).....	47
Encart 5 : les types d'intersections des coniques et la décomposition polynomiale du déterminant.....	48
Encart 6 : les types d'intersection des quadriques.....	53
Encart 7 : les perceptions du mémoire de Cayley (1858) à la fin du XIX <sup>e</sup> siècle et l'écriture d'une histoire.....	54
Encart 8 : Quelques éléments biographiques sur Eduard Weyr.....	56
Encart 9 : plan d'une démonstration contemporaine du théorème de Jordan.....	57
Encart 10 : le rôle de la représentation matricielle dans la formulation contemporaine de la relation mathématique entre diviseurs élémentaires et forme de Jordan.....	58
Encart 11 : comparaison du calcul des matrices de Weyr et de la théorie des matrices de Cayley.....	59
Encart 12 : comparaisons entre Weyr et Cayley.....	60
Encart 13 : la forme canonique de Jordan en 1870.....	62
Encart 14 : Pincherle et la recherche de sous variétés d'un espace vectoriel invariants par rapport à une homographie A.....	63
Encart 15 : extraits de différents théorèmes énoncés dans des réseaux de recherches distincts.....	63

## Introduction.

Un jeu de couleurs porté sur un texte mathématique contemporain sollicite différents moments de l'histoire sur une période longue et fait apparaître une interrogation historique.

La **matrice** suivante

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & -2 & 0 \\ -2 & 2 & 1 \end{pmatrix}$$

peut non seulement représenter un **tableau** de nombres, mais aussi un **déterminant**, une **famille de vecteurs**, une **application linéaire**  $E \rightarrow F$ , un **endomorphisme**  $E \rightarrow E$ , un **système d'équations différentielles linéaires**, une **forme bilinéaire**. D'une part, le mathématicien contemporain peut interpréter de *différentes* manières une *même* représentation matricielle. D'autre part un *même* objet peut se représenter par des matrices *différentes*. Par exemple, la matrice  $A$  et la matrice  $B$  ci-dessous représentent une même application linéaire  $E \rightarrow E$  dans des bases différentes :

$$B = \begin{pmatrix} \frac{3}{14} & -\frac{5}{14} & -\frac{43}{7} \\ \frac{3}{2} & \frac{5}{2} & 6 \\ -\frac{13}{14} & -\frac{11}{14} & -\frac{26}{7} \end{pmatrix}$$

- **Matrice.**

**Sylvester, 1851. Géométrie.** "I have in a previous paper defined a « Matrix » as a rectangular array of terms, out of which different systems of determinants may be engendered, as from the womb of a common parent [...]."

**Cayley, 1858. Théorie des matrices.** " [...] a set of quantities arranged in the form of a square, e.g.

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$$

is said to be a matrix [...]."

**Weyr, 1890. Théorie des formes bilinéaires.** "Eine Matrix  $n^{\text{ter}}$  Ordnung  $\{a_{hk}\}$  oder  $A$  auf ein System von  $n$  Werten  $x_1, \dots, x_n$ , oder kürzer  $(x)$  applicieren, heißt ein Werthsystem  $y_1, \dots, y_n$  mittelst der  $n$  linearen Gleichungen

$$y_h = a_{h1}x_1 + \dots + a_{hn}x_n \quad (h=1, 2, \dots, n)$$

ableiten. Diese  $n$  Gleichungen mögen durch die symbolische Gleichung

$$(y) = A(x)$$

ausgedrückt werden. »

- **Tableau.**

**Poincaré, 1884. Groupes de Lie.** "Ecrivons le Tableau à double entrée des coefficients d'une substitution quelconque de ce groupe [...]. Dans ce Tableau, séparons par des traits verticaux les  $\alpha$  premières colonnes, puis les  $\beta$  suivantes etc., puis les  $x$  dernières. Séparons de même par des traits horizontaux les  $\alpha$  premières lignes, puis les  $\beta$  suivantes, etc., puis les  $x$  dernières. Nous avons partagé nos coefficients en  $p^2$  systèmes. Si l'on choisit convenablement les  $n$  paramètres arbitraires en fonction desquels tous les coefficients du groupe s'expriment linéairement, un quelconque d'entre eux ne pourra entrer que dans les coefficients d'un seul des  $p^2$  systèmes. »

**Châtelet, 1911. Théorie des nombres.** « [...], il m'a semblé commode d'introduire la notion de tableaux ou systèmes linéaires. De tels symboles ont en effet l'avantage de représenter des êtres mathématiques assez divers : systèmes de formes linéaires, forme bilinéaire, forme décomposable, substitution linéaire [...]. »

- **Application linéaire.**

**Jordan, 1870, Théorie des substitutions.** "Réduction d'une substitution linéaire à sa forme canonique simple. »

**Pincherle, 1899, Equations différentielles.** "Décomposition d'un espace vectoriel en sous variétés invariantes par l'opération d'une homographie  $A$ ."

- **Etc...**

Que l'on s'attache au terme de tableau, de **forme bilinéaire**, d'**endomorphisme**, d'**équation différentielle**, *différents* auteurs, *différents* époques, *différents* contextes culturels sont mobilisés par la lecture d'un *même* texte. L'histoire de la notion de matrice débute-t-elle en 1858, date de parution d'un célèbre mémoire de Cayley ? Doit-on s'attacher au *calcul des Tableaux* très employé en France, au XIX<sup>e</sup> siècle, de Cauchy à Poincaré ? Faut-il remonter plus loin dans le temps et étudier les systèmes linéaires d'équations différentielles résolus par

des savants du XVIII<sup>e</sup> siècle comme d'Alembert, Lagrange et Laplace ? *Y a-t-il une ou plusieurs notions de matrice dans l'histoire ?* La question de l'identité du concept de matrice donne à cet article sa trame principale <sup>(1)</sup>.

Au cours des années trente du XX<sup>e</sup> siècle, la notion de matrice est devenue un élément de base dans l'architecture du savoir algébrique. L'acquisition de ce statut élémentaire au sein d'une théorie, l'algèbre linéaire, a donné aux matrices *une identité forte* et, dans le même temps, a écrasé *la pluralité de leur histoire*. Porter un regard sur des périodes antérieures à l'algèbre linéaire des années 1930 nécessitera de distinguer, dans l'histoire d'une *même* notion, des *concepts* et des *pratiques différentes* qui ne sont pas étrangères à la variété de représentations qu'offre au mathématicien la notion de matrice. Nous verrons toutes les différences entre la notion originelle de James Joseph Sylvester (1851) d'une matrice comme *mère des mineurs d'un déterminant*, les *lois du calcul des matrices* d'Arthur Cayley (1858) ou encore les *procédés de décomposition matricielle* d'Eduard Weyr (1885).

L'histoire de la notion de matrice ne peut pas se réduire à *l'histoire d'une théorie avant l'algèbre linéaire des années trente du XX<sup>e</sup> siècle*, elle doit plutôt s'envisager *comme le filage en une tresse de fils reliant des époques et des contextes variés* <sup>(2)</sup>. Plutôt que de proposer une vision schématique de la longue histoire séparant 1850 et 1930, cet article développe quelques moments particuliers afin de faire apparaître des *implicites* du savoir mathématique qui tiennent à des *modes de pensées* et à des *pratiques* indissociables d'un *contexte culturel daté*. Nous verrons que l'histoire des matrices permet de dévoiler des *aspects culturels* des mathématiques *antérieurs aux théories structurelles et unificatrices* comme l'algèbre linéaire des années 1930 <sup>(3)</sup>. La richesse de l'histoire des matrices provient souvent de ce qui échappe à une description mathématique contemporaine et, pour cette raison, des extraits de textes originaux accompagnent le texte principal sous forme d'encarts <sup>(4)</sup>.

---

<sup>1</sup> Cet article est issu de recherches effectuées dans le cadre d'une thèse de doctorat, intitulée *Histoire du théorème de Jordan de la décomposition matricielle* et menée sous la direction de Jean Dhombres au centre Alexandre Koyré - Ecole des Hautes Etudes en Sciences Sociales [Brechenmacher, 2006].

<sup>2</sup> La métaphore de la tresse renvoie plus généralement à une position sur la manière de restituer la dynamique réelle des savoirs par la multiplicité de leurs origines qui revient à ancrer une notion mathématique dans l'histoire en posant la question de son identité. Dans son ouvrage intitulé *Un théorème de Fermat et ses lecteurs*, Catherine Goldstein a montré la pertinence de la question de l'identité pour décrire des évolutions qui ne relèvent pas simplement d'une activité mathématique de recherche de nouveaux résultats ou de meilleures preuves, l'identité y est présentée : "comme un problème et non comme une tautologie : elle témoigne de pratiques et, à travers elles, de la manière dont est estimée l'innovation" [Goldstein, 1995, 16]. Voir également la réflexion sur la multiplicité des origines des savoirs mathématiques développée par Jean Dhombres [Dhombres, 2002].

<sup>3</sup> L'étude des fluctuations d'élaborations mathématiques sur une longue période présentée dans l'ouvrage collectif sur l'histoire des fractions coordonné par P. Benoit, K. Chemla et J. Ritter [Benoit et al., 1992] et le regard de C. Goldstein sur les relations entre analyse diophantienne et descente infinie [Goldstein, 1993] ont été des sources d'inspiration pour ce travail qui s'inscrit dans une entreprise plus vaste visant une histoire plurielle de l'algèbre linéaire.

<sup>4</sup> Comme l'écrivait Hourya Sinaceur dans son histoire du théorème de Sturm, "pour une histoire réfléchie il est encore plus important que pour un memento de résultats de revenir, par delà les traditions didactiques, aux mémoires originaux. On y apprend toutes les identités que le "progrès" efface : identité d'un contexte, d'un objectif, d'une perspective, d'un langage, sans parler de tout ce qui reste implicite sans manquer d'être là" [Sinaceur, 1991, 21].

# I. La théorie des matrices canoniques des années trente du XXe siècle.

Dans un premier temps, porter notre regard sur la période qui voit la notion de matrice acquérir un *caractère universel* au sein d'une *théorie internationale* va nous permettre de dégager les principales problématiques et la méthodologie de cet article.

## 1. La nouveauté de la théorie des matrices des années 1920-1930.

On publie beaucoup sur les matrices dans les années 1920-1930. Des *formes de représentations imaginées* envahissent des textes mathématiques publiés dans toutes les langues, dans le domaine de la recherche comme dans celui de l'enseignement. Comment une notion, dont Cayley présentait déjà une théorie en 1858, peut-elle porter une *nouveauté* dans les années folles ? Regardons les arguments développés dans un ouvrage publié en 1932 par H.W. Turnbull et A.C. Aitken, *An Introduction to the Theory of Canonical Matrices* ([encart 1](#)). Selon l'introduction du traité, la *nouveauté* de la théorie des "matrices canoniques" est dans l'utilisation des "propriétés" de l'"idiome matriciel" par opposition à une théorie *classique*, la *théorie des formes bilinéaires*, attachée à de grands noms de la fin du XIX<sup>e</sup> siècle comme Weierstrass, Kronecker et surtout Frobenius.

$$\begin{array}{l} \text{Notation matricielle : } \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \end{array} \qquad \begin{array}{l} \text{Notation des formes bilinéaires : } \sum_{i,j} a_{ij} x_i y_j \end{array}$$

Selon Turnbull et Aitken (T&A), le recours à la notation matricielle marque une *rupture* par rapport à une *pratique traditionnelle* des traités d'algèbre des générations précédentes <sup>(5)</sup>. Deux arguments sont développés pour justifier cette évolution. D'abord, la notation matricielle est présentée comme *efficace* car permettant des énoncés de "théorèmes généraux" en un "minimum de place". Surtout, et comme tous les auteurs des années 1920-1930, T&A mettent en avant les *valeurs pédagogiques* d'une représentation présentée comme "simple". Dans l'introduction de son traité sur la théorie des groupes publié en 1916, Blichfeldt encourageait déjà l'étudiant à bénéficier des "avantages" procurés par l'"image mentale" de la "forme matricielle" par opposition à la "représentation linéaire".

From the outset the student is urged to work with the matrix form of a linear representation. The practice thus gained is of great advantage [...]; in particular the more difficult sections of Chapter XIII [group representation] will be mastered readily if the student has a clear mental image of the matrix form of the regular groups [...]. [Blichfeldt, 1916, vii].

Traduction, F.B. :

Nous encourageons l'étudiant à travailler dès le départ avec la forme matricielle d'une représentation linéaire. Il y gagnera une pratique très avantageuse et les parties plus difficiles du chapitre XIII (sur la représentation des groupes) seront maîtrisées rapidement si l'étudiant a une image mentale de la forme matricielle des groupes classiques.

Si, au début du siècle, les partisans de la représentation en tableau, comme Autonne et Chatelet en France, Cullis en Grande Bretagne, étaient encore rares, l'idée selon laquelle la

<sup>5</sup> Cette popularité des matrices n'est que temporaire. Après 1940, la prise en compte de la dimension infinie dans la théorie des opérateurs donnera une importance moindre à la représentation matricielle.

maîtrise de la *représentation matricielle* permettrait d'assimiler plus "simplement" des "théorèmes généraux" de différentes théories est à la base des *organisations didactiques de nombreux traités d'algèbre publiés dans les années 1930*. Comme le montre l'introduction de l'ouvrage de T&A, la représentation matricielle ne peut être dissociée de l'objet même de la nouvelle organisation théorique développée dans le traité :

La théorie des matrices canoniques a pour objet l'investigation systématique des types de transformations qui réduisent les matrices à leur forme la plus simple et la plus pratique.  
[Turnbull et Aitken, 1932, 2, traduction F.B.].

La théorie porte sur les *formes de représentation* des matrices : il s'agit de *réduire* les matrices à leurs *formes les plus simples* ou *formes canoniques*, c'est-à-dire d'élaborer des *pratiques* de "transformations", de "réductions" des "formes", normées par un critère de "simplicité". Les résultats principaux énoncent des formes canoniques pour les relations *d'équivalence*, de *similitude* ou de *congruence* des matrices. Pour la relation de similitude, *deux formes canoniques* sont associées à un *unique théorème de décomposition matricielle* dont un énoncé est donné en [encart 2](#) (dans la citation suivante, le symbole « . » représente la valeur 0) :

$$A = \begin{bmatrix} \lambda & 1 & . \\ . & \lambda & 1 \\ . & . & \lambda \end{bmatrix}, B = \begin{bmatrix} . & 1 & . \\ . & . & 1 \\ \lambda^3 & -3\lambda^2 & 3\lambda \end{bmatrix}, (13)$$

we may prove in the following manner that there exists a matrix  $H$  such that  
 $HAH^{-1}=B$ .

[Turnbull et Aitken, 1932, 2].

D'une part la matrice  $A$ , "forme de Jordan", est la *forme la plus simple* et donne une *décomposition maximale* ; d'autre part la matrice  $B$ , qualifiée de "forme rationnelle", est obtenue par des *procédés effectifs* <sup>(6)</sup>. L'identité d'un *unique* théorème de décomposition en *deux* formes canoniques se décline sous une forme mathématique : on passe d'une forme canonique à l'autre par une méthode spécifique qui se caractérise comme une *combinatoire* sur la *représentation*, sur la *forme des matrices*. Les premiers chapitres du traité de T&A développent une *pratique algébrique de décomposition* basée sur un *caractère opératoire* donné aux "sous-matrices" que l'on combine les unes avec les autres comme on le voit dans les extraits suivants :

#### Matrices partitioned into Submatrices.

It is convenient to extend the use of the fundamental laws of combination for matrices to the case where a matrix is regarded as constructed not so much from elements as from submatrices, or minor matrices, of elements. For example, the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

can be written

<sup>6</sup> C'est-à-dire que l'obtention de la première forme exige de pouvoir extraire les racines d'une équation algébrique générale et donc de se placer dans un corps algébriquement clos. La seconde forme, pour laquelle on utilise aujourd'hui la désignation donnée par Krull de "matrice compagnon", peut s'obtenir pour toute matrice à coefficients dans un anneau principal. En 1874, la *simplicité* de la première forme canonique et les *procédés effectifs* attachés à la seconde étaient opposés lors d'une vive querelle entre Camille Jordan et Léopold Kronecker, voir à ce sujet [Brechenmacher, 2006].



$$A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

where

$$P = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, Q = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, R = [7, 8], S = [9].$$

Comme l'illustrent les extraits suivants et la démonstration du théorème de Jordan donnée en [encart 3](#), un *caractère opératoire* est conféré à la représentation matricielle que l'on *décompose* en cultivant l'analogie avec les figures de la *géométrie* (on parle ainsi de rectangle, triangle, diagonale). La méthode matricielle se construit en articulant des pratiques *combinatoires* d'extractions de sous matrices, des *décompositions polynomiales*, un *calcul symbolique* des puissances de matrices, une *arithmétique* des lignes et des colonnes et le point de vue *vectériel de décomposition d'un espace* en *sous espaces stables* par l'action d'une transformation linéaire.

Here the diagonal submatrices  $P$  and  $S$  are **square**, and the partitioning is **diagonally symmetrical**. In the general case there may evidently be  $n$  or fewer partitions **row-wise** or **column-wise**. Let  $B$  be a second **square** matrix of the third order similarly partitioned:

$$B = \begin{bmatrix} P_1 & Q_1 \\ R_1 & S_1 \end{bmatrix} = \begin{bmatrix} 2 & . & 1 \\ 3 & 1 & 2 \\ 1 & 2 & . \end{bmatrix}$$

then by **addition** and **multiplication** we have

$$A+B = \begin{bmatrix} P+P_1 & Q+Q_1 \\ R+R_1 & S+S_1 \end{bmatrix}, AB = \begin{bmatrix} PP_1+QR_1 & PQ_1+QR_1 \\ RP_1+SR_1 & RQ_1+SS_1 \end{bmatrix}, \text{ as may readily be verified. In each}$$

case the resulting matrix is of the same order, and is partitioned in the same way, as the original **matrix factors**. For example, in  $AB$  the first element,  $PP_1+QR_1$ , stands for a **square** submatrix of **two rows** and **columns**: and this is possible, since, by definition, both **products**  $PP_1$  and  $QR_1$  consist of two **rows** and two **columns**. [...]. Thus

$$PQ_1 + QR_1 = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} [0] = \begin{bmatrix} 5 \\ 14 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 14 \end{bmatrix},$$

giving the proper **rectangular shape** for the **upper right** hand minor.

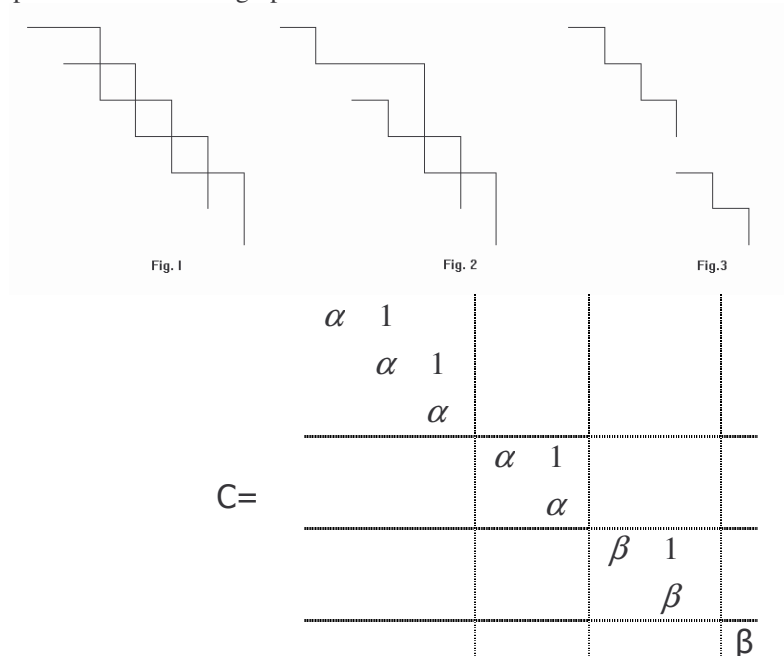
[Turnbull et Aiken, 5-6]

Les différentes démonstrations du théorème de Jordan, les applications à différents problèmes comme la recherche des matrices commutant avec une matrice donnée sont autant d'exemples de l'efficacité d'une *pratique algébrique recourant à des formes*, des images que Picard désignait encore en 1910 comme des "dessins" et qui, dans les années trente, envahissent les textes mathématiques comme dans les extraits ci-dessous du traité de T&A.

$$\begin{array}{c} \mathbf{A} \end{array} \times \begin{array}{c} \mathbf{B'} \end{array} = \begin{array}{c} \mathbf{AB'} \end{array} ;$$

$$\begin{bmatrix} C_4(\alpha) & . \\ . & C_3(\alpha) \end{bmatrix} = \begin{bmatrix} \alpha & 1 & . & . \\ & \alpha & 1 & . \\ & & \alpha & 1 \\ & & & \alpha \\ & & & & \alpha & 1 & . \\ & & & & & \alpha & 1 \\ & & & & & & \alpha \end{bmatrix} .$$

[...] It is a help to form a staircase graphs of these chains as follows:



Dans les années 1920-1930, les procédés opératoires portés sur la représentation matricielle sont l'objet de communications dans les congrès et de publications dans toutes les langues. *Ces procédés fondent le caractère unificateur de la représentation matricielle à partir de pratiques algébriques auparavant distinctes*, ils font de la théorie des matrices une théorie internationale qui participe à la réorganisation du savoir algébrique menée à cette époque avec l'élaboration de *l'algèbre linéaire*.

## 2. Les problématiques liées à l'histoire des procédés matriciels.

Les valeurs pédagogiques et les procédés pratiques mis en avant par des auteurs comme T&A peuvent surprendre venant d'une époque qui voit l'émergence de structures basées sur des nouvelles notions souvent qualifiées *d'abstraites* et *unificatrices* comme les groupes, les modules ou les vecteurs. Comme nous l'avons vu au paragraphe précédent, le caractère *unificateur* de la théorie des matrices dans les années trente est pourtant fondé sur des *pratiques opératoires*, qui, loin d'être *abstraites*, sont associées à une *forme de représentation*



*imaginée*. Pour cette raison, le problème de l'histoire des procédés matriciels est souvent passé inaperçu de travaux historiques focalisés sur les notions abstraites et structurelles de l'algèbre linéaire <sup>(7)</sup>. La représentation matricielle a d'ailleurs souvent été utilisée au sein des discours historiques comme une représentation *inoffensive, naturelle, dénuée d'histoire* <sup>(8)</sup>. Son histoire n'en pose pas moins une véritable problématique. Comment comprendre que la théorie des matrices canoniques s'unifie en 1930 autour de procédés opératoires alors que de tels procédés étaient déjà présents dans un mémoire de Cayley en 1858 ? Pourquoi cet écart de 80 ans ? Si l'on peut soupçonner que la notion de matrice du XX<sup>e</sup> siècle n'est pas la *même* que celle de Cayley, il faut alors poser la question des *évolutions* dans l'identité de la notion de matrice en portant notre attention sur les *pratiques opératoires élaborées entre 1858 et 1930*. Mais comment choisir les textes et les auteurs, sur lesquels porter notre étude ? Avant les années trente, le caractère opératoire de la représentation en tableau n'était pas identifié comme un *problème* mathématique. Au début du siècle, par exemple, le rapport de Picard sur la thèse de Chatelet (1910), consacrée à *explicit* les pratiques du « calcul des tableaux », manifestait un grand scepticisme sur l'intérêt mathématique d'un travail de doctorat consacré à des « dessins ». Si le calcul des tableaux renvoyait déjà au temps de Chatelet à une tradition ancienne, forte des contributions de Cauchy, Hermite, Jordan ou Poincaré, les procédés associés restaient *implicites* <sup>(9)</sup>. *Comment aborder l'histoire des procédés opératoires associés aux matrices alors que la représentation qui les sous-tend n'est pas identifiée comme problématique avant la théorie des matrices canoniques des années trente ?*

### 3. La méthodologie des réseaux de textes.

La méthodologie qui a été mise en œuvre au sein de la thèse de doctorat dont cet article est issu repose sur l'établissement de réseaux de textes. Après avoir choisi un moment de référence, les années trente, une recherche bibliographique a été menée sur toutes les références des traités parus dans les années 1920-1930 <sup>(10)</sup>. Ce premier corpus de textes a ensuite été complété par épuisement systématique des références bibliographiques. L'examen du corpus général obtenu a permis de fixer une périodisation, allant de 1850 à 1930, dont un découpage plus précis a nécessité d'étudier la manière dont les textes et acteurs des périodes considérées s'organisent en réseaux. L'analyse des références bibliographiques permet de distinguer des *réseaux cohérents de textes*, essentiellement distincts les uns des autres et *ne*

<sup>7</sup> Dans le cadre de travaux historiques portant sur l'émergence de structures algébriques au début du XX<sup>e</sup> siècle (comme l'algèbre linéaire), l'accent est souvent porté sur le caractère conceptuel de la théorie des matrices. La représentation matricielle des années 1930 n'a bénéficié d'un éclairage historique que pour la distinction qu'elle réalise entre les concepts de matrices et de transformations linéaires, distinction permettant d'envisager les problèmes de représentations comme *relatifs au choix d'une base d'un espace vectoriel*.

Voir les passages consacrés à la notion de matrice par B.L. Van der Waerden [1977], T. Hawkins, [1975, 1977], J. Dieudonné [1978], K. Parshall [1988], I. Grattan-Guinness [1994] ou L. Corry [1996].

<sup>8</sup> Si la représentation matricielle paraît naturelle au mathématicien contemporain, son emploi au sein d'un texte d'histoire n'est pas inoffensif. Lorsque, par exemple, Thomas Hawkins [1975, 1977] employait des matrices pour décrire la méthode élaborée par Lagrange en 1766 pour la résolution des systèmes différentiels linéaires, des notions absentes du texte de Lagrange s'introduisaient subrepticement dans le discours historique. La représentation matricielle porte implicitement des idées anachroniques au texte de Lagrange comme l'idée géométrique de "matrice symétrique" ou la pratique consistant à "transformer" la matrice d'un système en une forme diagonale (la méthode de Lagrange n'était pas basée sur une transformation du système linéaire mais sur la méthode générale d'abaissement du degré d'une équation différentielle dont on connaît des solutions particulières).

D'autres exemples des implicites liés à la notation matricielle et une discussion détaillée de leurs implications sur l'historiographie de l'algèbre linéaire sont proposés dans [Brechenmacher, 2006, 280-487].

<sup>9</sup> Au sujet des travaux de Chatelet sur le calcul des tableaux, voir [Brechenmacher, 2006]. D'autres exemples de pratiques comme la combinatoire arithmétique des lignes et des colonnes de Kronecker (1880-1890) ou la décomposition des matrices de Weyr (1890) montrent que les procédés opératoires sur lesquels sera fondée la théorie des matrices des années trente ne s'élaborent pas par rapport un questionnement mathématique spécifique mais comme des pratiques développées en réponse à des problèmes divers et dans des cadres théoriques distincts

<sup>10</sup> La base de données utilisée regroupe les fonds du réseau national des bibliothèques de mathématiques, les archives de l'Ecole Polytechnique ainsi que des ressources numériques en lignes (*Bibliothèque Nationale de France, cellule math. doc., Göttinger Digitalisierungszentrum, universités de Cornell et du Michigan*).

*correspondant pas globalement à des théories.* Des graphes comme celui de l'[encart 4](#) permettent de représenter les liens entretenus par les différents textes d'un même réseau, ils montrent notamment l'existence de points de convergence, de nœuds, dans l'entremêlement des références bibliographiques. La méthodologie des réseaux permet de préciser la métaphore de la tresse appliquée à la notion contemporaine de matrice en introduction : les différentes pratiques qui s'entrelacent dans les procédés matriciels des années 1930 comme nous le voyons par exemple dans l'[encart 3](#) (analogies *géométriques*, pratiques *combinatoires*, *décompositions polynomiales*, *calcul symbolique*, *arithmétique*, point de vue *vectorel* etc.) sont élaborées souvent indépendamment les unes des autres et dans de réseaux distincts. Il faut donc décrire les conditions de ces élaborations avant de poser la question des communications, des convergences, c'est-à-dire de la manière dont des cultures locales se tressent et participent d'une histoire plurielle.

La seconde partie de cet article propose de suivre quelques fils de la tresse que constitue la représentation matricielle des années trente. Nous nous attacherons à l'examen d'un réseau particulier représenté en [encart 4](#) et dont les points de convergence mettent en avant les noms de Sylvester, Cayley, Weyr ou Frobenius.

## II. Une première origine : les travaux de Sylvester et Cayley des années 1850.

Les termes « matrice » et « mineurs » sont introduits par Sylvester en 1851 dans le cadre de travaux géométriques. Entre 1850 et 1851, Sylvester publie quatre mémoires consacrés au problème des *intersections de deux coniques ou quadriques* <sup>(11)</sup>. Ces publications successives permettent de suivre l'élaboration progressive d'une méthode qui se caractérise par une traduction de propriétés géométriques ou analytiques dans le cadre du *calcul des déterminants*. Les notions de « matrices » et « mineurs » apparaissent d'abord, pour reprendre les termes de l'auteur, comme des « résultats collatéraux » avant de devenir les principaux objets d'étude de Sylvester.

### 1. Origine des matrices comme mères des mineurs : les travaux de Sylvester [1850-1851] sur les intersections de deux coniques.

Le mémoire intitulé "On the intersections, contacts and other correlations of two conics expressed by indeterminate coordinates" [Sylvester, 1850a], publié en novembre 1850 dans le *Cambridge and Dublin Mathematical Journal*, porte sur l'étude des différents types d'intersection de deux coniques. La caractérisation des intersections de coniques avait déjà été traitée par Plücker en 1828 et l'originalité du travail de Sylvester réside surtout dans le recours au calcul des déterminants par opposition à la méthode analytique développée par les savants français de l'Ecole Polytechnique comme Hachette et Poisson (1802), Cauchy ou Biot (1826). Pour Sylvester, la méthode analytique traditionnelle est encombrée de la considération d'équations "arbitraires", au contraire du caractère intrinsèque du calcul des déterminants. Nous verrons que la *nature* des types de contacts de deux coniques  $U$  et  $V$ , « traduite dans le langage des déterminants », est liée à l'étude de la *multiplicité* des trois racines de l'équation  $|U + \mu V| = 0$  <sup>(12)</sup>. Mais il existe cinq cas d'intersections et seulement trois types de multiplicités des racines (3 racines simples, 1 double, 1 triple), les occurrences de racines doubles et triples nécessitent donc chacune l'examen de deux sous cas comme le détaille l'[encart 5](#). Afin de caractériser les cinq types d'intersection, Sylvester va comparer les *décompositions algébrique du polynôme  $|U + \mu V|$  aux décompositions du déterminant  $|U + \mu V|$  en "mineurs" extraits d'une matrice* <sup>(13)</sup>.

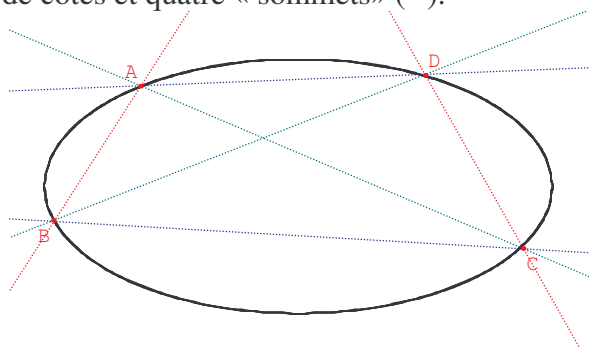
- **Le problème des intersections de deux coniques  $U$  et  $V$  et la multiplicité des racines de l'équation  $|U + \mu V| = 0$ .**

<sup>11</sup> Dans l'ordre chronologique, les quatre publications de Sylvester sont les suivantes : « On the intersections, contacts and other correlations of two conics expressed by indeterminate coordinates » [1850a], « Additions to the articles « On an new class of theorems » and on Pascal theorem » [1850b], « An enumeration of the contacts of lines and surfaces of the second order » [1851a], « On the relations between the minor determinants of linearly equivalent quadratic functions » [1851b].

<sup>12</sup> Sylvester emploie des coordonnées homogènes, la conique (projective)  $U$  est donc représentée par une forme quadratique à trois variables. En termes contemporains, si  $U$  et  $V$  sont les matrices dans la base canonique de  $\mathbb{E}^2$  des formes quadratiques définissant les deux coniques, l'équation  $|U + \mu V| = 0$  revient, si  $|V| \neq 0$  à  $|UV^{-1} + \mu I| = 0$  c'est-à-dire à une équation caractéristique. La nature du type de contact entre les deux coniques dépend alors de la multiplicité des valeurs propres de la matrice  $UV^{-1}$ .

<sup>13</sup> A la fin du XIX<sup>e</sup> siècle, les mathématiciens G. Darboux puis M. Noether présenteront les travaux Sylvester comme précurseurs de la notion de *diviseurs élémentaires* de couples de formes bilinéaires introduite par Weierstrass en 1868 ([encart 10](#)).

Dans son premier article de 1850, Sylvester reprend à son compte la caractérisation faite par Cayley de l'intersection de deux coniques  $U$  et  $V$  comme formant un « *quadrangle* » comportant trois paires de cotés et quatre « sommets »<sup>(14)</sup>.



Caractériser l'intersection des coniques revient alors à étudier la nature des sommets ou des paires de côtés du « *quadrangle* » (réels-imaginaires, distincts-confondus) :

If all the points of the quadrangle of intersection are real, the three vertices and the three pairs of sides are all real. If only two points of the quadrangle are real, one vertex and one of the three pairs of sides will be real; the other two vertices and two pairs of sides being imaginary. If all four points of the quadrangle are unreal, one pair of sides will be real and the other two pairs imaginary, as in the last case; but all the three vertices will remain real, as in the first case. [Sylvester, 1850a, 263].

Traduction, F.B.

Si tous les points du quadrangle d'intersection sont réels, les trois sommets et les trois paires de côtés sont tous réels. Si seulement deux points du quadrangle sont réels, l'un des sommets et l'une des trois paires de côtés seront réels ; les deux autres sommets et paires de côtés étant imaginaires. Si tous les quatre points du quadrangle sont non réels, une paire de côtés sera réelle et les deux autres paires imaginaires comme dans le cas précédent et, comme dans le premier cas, les trois sommets seront réels.

Toutes les coniques représentées par  $U + \mu V$ , avec  $\mu$  variable, se coupent en un même quadrangle. Le quadrangle lui-même correspond aux valeurs de  $\mu$  pour lesquelles la conique  $U + \mu V$  est une paire de droites, c'est-à-dire, en termes contemporains, une conique dégénérée du faisceau  $U + \mu V$ <sup>(15)</sup> :

Hence we have a direct and simple criterion for distinguishing the case of mixed intersection from intersection wholly real or wholly imaginary; namely, that the cubic equation of the roots of which coordinates of the vertices are real linear functions shall have a pair of imaginary roots. This is the sole and unequivocal condition required. The equation in question is, or ought to be, well known to be the determinant in respect to  $\xi, \eta, \zeta$  of  $\lambda U + \mu V$ . [Sylvester, 1850a, 264].

Traduction, F.B.

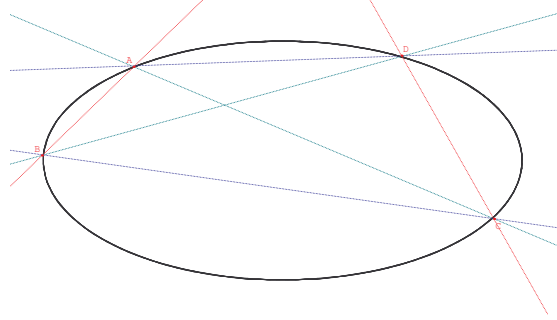
Nous avons donc un critère simple et direct pour distinguer le cas d'intersection mixte des cas d'intersections entièrement réelles ou entièrement imaginaires ; l'équation cubique dont les coordonnées des sommets sont des fonctions linéaires des racines doit avoir une paire de racines imaginaires. Cette condition est la seule requise. L'équation en question est bien connue comme étant le déterminant de  $\lambda U + \mu V$  par rapport à  $\xi, \eta, \zeta$  (les coordonnées projectives).

<sup>14</sup> Deux coniques projectives de  $P_2(\hat{E})$ ,  $C$  et  $C'$ , se coupent en quatre points, c'est un cas particulier du théorème de Bézout sur les courbes projectives complexes de degré  $m$  et  $n$  se coupant en  $mn$  points.

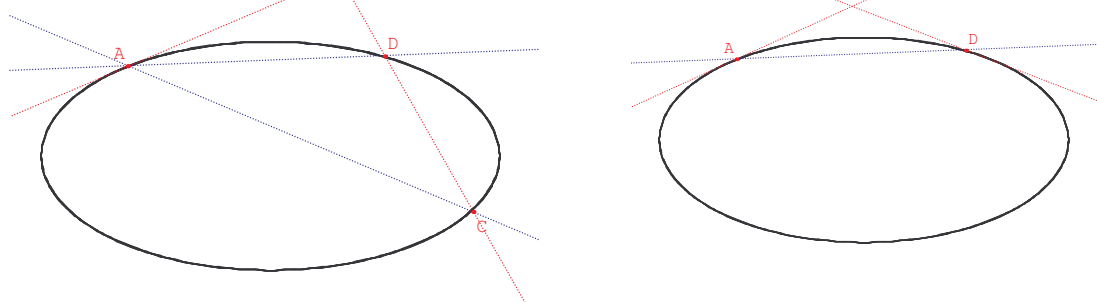
<sup>15</sup> Si l'on suppose  $C$  propre et que l'on appelle  $Q$  et  $Q'$  les formes quadratiques associées aux deux coniques, l'intersection des coniques  $C$  et  $C'$  est celle de  $C$  avec n'importe quelle conique du faisceau défini par  $\lambda Q(x,y,z) + \mu Q'(x,y,z) = 0$ . Il s'agit en particulier de l'intersection de  $C$  avec une conique dégénérée  $C''$  du faisceau (paire de droites), la conique  $C''$  correspondant à des valeurs de  $\lambda$  et  $\mu$  telles que

$$D(\lambda) = \det(\lambda Q(x,y,z) + \mu Q'(x,y,z)) = 0.$$

La caractérisation du contact des coniques est ainsi ramenée à l'examen des racines de l'équation  $|U+\mu V|=0$ . L'occurrence de trois racines distinctes traduit l'existence d'un quadrangle formé par trois paires de droites distinctes.



Mais le seul degré de multiplicité des racines du déterminant de  $U+\mu V$  s'avère un critère insuffisant pour distinguer deux types de contact de second degré correspondant à l'occurrence d'une ou deux 'tangentes doubles' à la conique  $U$  ([encart 5](#)).



The classification of contacts between two conics may be stated as follows :

Simple contact = one case.

Second degree contact = 2 cases: common curvature or double contact.

Third degree contact = one case, namely, contact in four consecutive points

These four cases of course correspond to the several suppositions of there being two equal roots, three equal roots, two pairs of equal roots, or four equal roots in the biquadratic equation obtained between two variables by elimination performed in any manner between the given equations in the two conics. [Sylvester, 1850a, 266].

Traduction. F.B.

L'énoncé suivant donne la classification des contacts entre deux coniques :

Contact simple = 1 cas.

Contact de second degré = 2 cas : courbure commune ou contact double.

Contact du troisième degré = 1 cas correspondant au contact en quatre points consécutifs.

Ces quatre cas correspondent aux différentes occurrences d'égalités entre les racines, selon qu'il y ait deux racines égales, trois racines égales, deux paires de racines égales ou quatre racines égales [...].

Sylvester traduit l'occurrence de racines doubles dans le langage des déterminants. Des facteurs communs doivent apparaître dans le développement polynomial du déterminant:

$$\begin{vmatrix} \frac{d^2W}{d\xi^2} & \frac{d^2W}{d\xi d\eta} & \frac{d^2W}{d\xi d\zeta} & p \\ \frac{d^2W}{d\eta d\xi} & \frac{d^2W}{d\eta^2} & \frac{d^2W}{d\eta d\zeta} & q \\ \frac{d^2W}{d\zeta d\xi} & \frac{d^2W}{d\zeta d\eta} & \frac{d^2W}{d\zeta^2} & r \\ p & q & r & 0 \end{vmatrix}$$

$$= Ap^2 + Bq^2 + Cr^2 + 2Fqr + 2Grp + 2Hpq,$$

where all the coefficients are quadratic functions of  $\mu$ , and make

$$A=0, B=0, C=0, F=0, G=0, H=0,$$

each of these six equations in  $\mu$  will have one and the same root in common.

[...] I must observe that besides the equations involved in the condition that  $A, B, C, F, G, H$ , or which is the same thing, that any three of them shall all have a factor in common, we must have  $\Delta(U+\lambda V)$  containing the square of such common factor. [Sylvester, 1850a, 271].

Traduction. F.B.

[...] Je dois observer qu'en raison des équations caractérisant la condition selon laquelle  $A, B, C, F, G, H$ , doivent avoir un facteur commun, le carré d'un tel facteur commun doit être contenu dans  $\Delta(U+\lambda V)$ .

Il y a donc une relation entre les types d'intersections des coniques et les types de facteurs communs intervenant dans le développement polynomial du déterminant  $|U+\mu V|$  ([encart 5](#)). L'étude des différentes décompositions polynomiales de la « caractéristique »  $U+\mu V$  va alors se substituer aux préoccupations géométriques initiales et devenir le principal objet d'étude de Sylvester. C'est dans ce cadre que Sylvester introduit la notion de "mineurs" du déterminant  $|U+\mu V|$  comme le montre l'extrait suivant dans lequel l'auteur fait le récit de l'évolution de ses travaux sur les coniques (la notation  $(U+\lambda V)$  désigne le déterminant de  $U+\lambda V$ ) :

Wishing to subject it to analytical test, he found it necessary to obtain the condensed forms which serve to characterize the confluent contact of conics. In this way he became aware of the great utility of these condensed forms, and of the desideratum to be supplied in obtaining a complete list of them applicable to all varieties of contact. The happy thought then occurred to him of inverting the process which he had applied in the treatment of the contacts of conics, in the November Number of the Cambridge and Dublin Mathematical Journal ; for whereas the nature of the contacts was there assumed and translated into the language of determinants, he soon discovered that it was the more easy and secure course to assume the relations of every possible immutable kind that could exist between the complete and minor determinants corresponding to characteristics, by aid of these relations to construct the characteristics, and from the characteristics so obtained, determine the geometrical character of each resulting species of contact. [...] It is easily seen that to every relation of equality between the roots of the determinant of  $U+\lambda V$  must correspond a particular species of contact between the loci which  $U$  and  $V$  characterize. But we should make a great mistake were we to suppose that every such relation of equality between the roots of the determinant of  $U+\lambda V$  must correspond a particular species of contact; for instance, the characteristics of  $U$  and  $V$  of two conics are functions of three letters, and  $(U+\lambda V)$  will be a cubic function of  $\lambda$ . Such a function may have two roots, or all its roots equal: this would seem to give but two species of contact, whereas we well know that there are no less than four species of contact possible between two conics. Accordingly we shall find, that, in order to determine the distinctive characters of each species of contact, we must look beyond the complete determinant, and examine into the relations of the several systems of minor determinants that can be formed from  $U+\lambda V$  [...]. If  $U$  and  $V$  are characteristics of the two loci whose contacts are to be considered,  $U+\lambda V$  will be the function, the properties of whose complete determinant, and of minor systems of determinants belonging to it, will serve to specify the nature of the contact. [Sylvester, 1851a, 295].

Traduction. F.B.

Désirant soumettre (la caractérisation des contacts de coniques) à un questionnement analytique, il (Sylvester) pensa nécessaire d'obtenir les formes condensées permettant de caractériser le cas de « confluent contact » des coniques (c'est-à-dire la décomposition polynomiale de la caractéristique, voir [encart 5](#)). Il réalisa alors la grande utilité de telles formes condensées et conçut le désir d'en obtenir une liste complète applicable à tous les cas de contact. Il lui vint alors l'idée heureuse d'inverser le procédé qu'il avait mis en œuvre pour son étude des contacts des coniques [...]; en effet, alors que, dans un premier temps, la nature des contacts avait été envisagée par une traduction dans le langage des déterminants, il découvrit qu'il était en fait plus rapide et plus sûr de commencer par étudier les différentes relations possibles entre le déterminant correspondant à la caractéristique et ses mineurs puis de construire les caractéristiques afin de déterminer finalement le caractère géométrique des espèces de contacts correspondantes aux caractéristiques obtenues [...]. Il est facile de voir qu'à toute occurrence d'égalité entre les racines du déterminant de  $U+\lambda V$  doivent correspondre des espèces particulières de contacts entre les lieux caractérisés par  $U$  et  $V$ . Mais nous ferions une grave erreur si nous supposions qu'à chaque relation d'égalité entre les racines correspond une espèce particulière de contact ; par exemple, les caractéristiques de deux coniques  $U$  et  $V$  sont des fonctions de trois lettres et  $(U+\lambda V)$  est donc une fonction cubique de  $\lambda$ . Une telle fonction peut avoir deux ou trois de ses racines égales, ce qui



ne donnerait que deux espèces de contacts entre deux coniques alors qu'il est bien connu qu'il n'y a pas moins de quatre espèces de contacts possibles. Afin de déterminer le caractère distinctif de chaque espèce de contact, nous devons par conséquent conclure qu'il est nécessaire de regarder au-delà du déterminant et en examinant les relations entre les différents systèmes de mineurs susceptibles d'être formés à partir de  $U + \lambda V$  [...]. Si  $U$  et  $V$  sont les caractéristiques des lieux dont les contacts sont étudiés, la nature du contact sera caractérisée par les propriétés des systèmes de mineurs contenus dans le déterminant de la fonction  $U + \lambda V$ .

Afin de caractériser les différents cas d'intersections correspondant à une même décomposition polynomiale de la caractéristique  $|U + \mu V|$ , il faut aller « au-delà du déterminant » et considérer les différents « systèmes de mineurs » que l'on peut en extraire. Comme on peut le voir en [encart 5](#), Sylvester associe à chaque type d'intersection la décomposition en "systèmes de mineurs" de l'expression  $|U + \lambda V|$ <sup>16</sup>. Le cas de « confluent contact » se caractérise par exemple par l'occurrence d'une racine triple et de « l'annulation simultanée des mineurs » :

If the determinant of this function ( $U + \mu V$ ) have two equal roots, the conics simply touch; if it have three equal roots, the conics have a single contact of a higher order, that is, the same curvature; if its six first minors become zero simultaneously for the same value of  $\mu$ , which makes all these first minors zero, be at the same time not merely a double root (as of analytical necessity it always must be) but a triple root of  $(U + \mu V) = 0$ . Then the conics have a single contact of the highest possible order short of absolute coincidence, that is, they meet in four consecutive points. [Sylvester, 1851a, 296].

Traduction. F.B.

Si le déterminant de cette fonction ( $U + \mu V$ ) a deux racines égales, les coniques se touchent simplement, dans le cas où il a trois racines égales, les coniques ont un contact de plus grand ordre, c'est-à-dire la même courbure : si tous les six premiers mineurs s'annulent simultanément pour la même valeur de  $\mu$ ,  $(U + \mu V) = 0$ , on aura alors une racine triple. Le contact des coniques sera alors du plus grand ordre qu'il est possible d'obtenir, en dehors du cas où elles seraient confondues, les coniques se rencontreront donc en quatre points consécutifs.

- **Généralisation de la méthode de Sylvester et introduction des matrices.**

Dans deux mémoires publiés en 1851, Sylvester *généralise* ses travaux aux intersections de quadriques et, plus généralement, à des formes quadratiques  $U$  et  $V$  portant sur «  $n$  lettres ». L'extraction effective des mineurs d'un déterminant d'ordre  $n$  s'appuie alors sur une représentation en tableau rectangulaire que Sylvester dénomme la "matrice" des "mineurs" :

[...] we must commence, not with a square, but with an oblong arrangement of terms consisting, suppose, of  $m$  lines and  $n$  columns. This will not in itself represent a determinant, but is, as it were, a Matrix out of which we may form various systems of determinants by fixing upon a number  $p$  and selecting at will  $p$  lines and  $p$  columns, the square corresponding to which we may be termed determinants of the  $p^{\text{th}}$  order. [Sylvester, 1851a, 296].

Traduction. F.B.

[...] Nous devons commencer avec, non pas un carré, mais un arrangement oblong de termes constitué, par exemple, de  $m$  lignes et  $n$  colonnes. Cet arrangement lui-même ne représente pas un déterminant mais une Matrice à partir de laquelle nous pouvons former différents systèmes de déterminants en fixant un nombre  $p$  and et en sélectionnant à volonté  $p$  lignes and  $p$  colonnes, les carrés correspondants pouvant être dénommés déterminants du  $p^{\text{e}}$  ordre. [Sylvester, 1851a, 296].

<sup>16</sup> En termes contemporains, la nature du contact de deux coniques est caractérisée par le système d'invariants formé du déterminant  $D$  et des plus grands communs diviseurs de ses sous-déterminants  $D_1, D_2, D_3, \dots$ . Le type d'intersection de deux coniques étant, comme nous l'avons vu, caractérisé par le type de coniques dégénérées se trouvant dans le faisceau, il dépend de la nature des racines (valeurs propres) de  $\det(\lambda Q + \mu Q') = 0$ . La multiplicité de ces racines permet de distinguer entre contacts de premier, second et troisième ordre. Il y a cependant deux types de points de contact pour les second et troisième ordres et, pour les différencier, il est nécessaire d'étudier les pgcd des mineurs successifs du déterminant  $D(\lambda)$  de  $\lambda Q(x, y, z) + \mu Q'(x, y, z)$ . Ces pgcd  $D_i(\lambda)$  sont les *facteurs invariants* de la matrice polynomiale  $\lambda Q + \mu Q'$ .



Le mémoire, « On the relations between the minor determinants of linearly equivalent quadratic functions » [1851b], est consacré à l'énoncé des propriétés des mineurs comme l'invariance par transformation linéaire. C'est dans ce cadre qu'est définie explicitement la notion de "matrice" comme *mère des mineurs* d'un déterminant <sup>(17)</sup>:

I have in previous papers defined a « Matrix » as a rectangular array of terms, out of which different systems of determinants may be engendered, as from the womb of a common parent ; these cognate determinants being by no means isolated in their relations to one another, but subject to certain simple laws of mutual dependence and simultaneous deperition. The condensed representation of any such matrix, according to my improved Vandermondian notation, will be

$$\left\{ \begin{matrix} a_1, a_2, \dots a_n \\ \alpha_1, \alpha_2 \dots \alpha_n \end{matrix} \right\}.$$

[...] It is wonderful that a theory so purely analytical should originate in a geometrical speculation. My friend M. Hermite has pointed out to me, that some faint indications of the same theory may be found in the *Recherches Arithmétiques* of Gauss. The notation which I have employed for determinants is very similar to that of Vandermonde, with which I have become acquainted since writing the above, in Mr Spottiswoode's valuable treatise *On the elementary Theorems of Determinants*. Vandermonde was evidently on the right road. I do not hesitate to affirm, that the superiority of his and my notation over that in use in the ordinary methods is as great and almost as important to the progress of analysis, as the superiority of the notation of the differential calculus over that of the fluxional system. For what is the theory of determinants? It is an algebra upon algebra; a calculus which enables us to combine and foretell the results of algebraical operations, in the same way as algebra itself enables us to dispense with the performance of the special operations of arithmetic. All analysis must ultimately clothe itself under this form. [Sylvester, 1851b, 296].

Traduction. F.B.

J'ai défini dans une publication précédente une "Matrice" comme une suite rectangulaire de termes de laquelle différents systèmes de déterminants peuvent être engendrés, comme de l'utérus d'un parent commun ; ces déterminants apparentés sont loin d'être isolés dans leurs relations l'un à l'autre, ils sont sujets à certaines lois simples de dépendance mutuelle et d'annulation simultanée. La représentation condensée d'une telle matrice, reprenant la notation que j'ai développée à partir de celle de Vandermonde, sera

$$\left\{ \begin{matrix} a_1, a_2, \dots a_n \\ \alpha_1, \alpha_2 \dots \alpha_n \end{matrix} \right\}$$

[...] Il est fantastique qu'une théorie d'une telle pureté analytique trouve son origine dans des spéculations géométriques. Selon mon ami Hermite quelques indications pourraient laisser supposer qu'une théorie semblable se trouverait dans les *Recherches Arithmétiques* de Gauss. Ma notation est similaire à celle de Vandermonde et j'en suis devenu familier la publication de mon précédent mémoire [celui de novembre 1850] à l'aide du remarquable traité de Spottiswoode, *Théorèmes élémentaires sur les déterminants*. Vandermonde était dans la bonne voie et je n'hésite pas à affirmer la supériorité de notre notation sur celle employée par les méthodes ordinaires, cette notation a une importance pour les progrès de l'analyse comparable à la supériorité de la notation du calcul différentiel sur le système des fluxions <sup>(18)</sup>. Car en effet, qu'est ce que le calcul du déterminant sinon une algèbre de l'algèbre, un calcul qui permet de combiner et prévoir les résultats d'opérations algébriques de manière similaire à l'algèbre elle-même qui nous dispense des opérations particulières de l'arithmétique.

Comme conclusion partielle, nous pouvons remarquer que *la notion de matrice est, à son origine chez Sylvester, associée aux méthodes du calcul du déterminant et plus particulièrement à une pratique spécifique qui articule des extractions de mineurs à une*

<sup>17</sup> Observer, dans cette citation, les références à Hermite et Gauss qui permettent de préciser le réseau arithmétique auquel se rattachent les travaux de Sylvester sur les formes quadratiques. A cette époque, Hermite, Cayley et Sylvester, ont une correspondance suivie. Hermite est notamment à l'origine de l'influence des *Disquisitiones Arithmeticae* de Gauss [1801] sur les travaux de mathématiciens anglais comme Sylvester, Cayley et Smith.

<sup>18</sup> Nous reviendrons plus loin sur cette allusion à la supériorité du calcul différentiel de Leibniz sur la notation des fluxions de Newton. Cette remarque de Sylvester manifeste un héritage de l'école algébrique anglaise de la première moitié du XIX<sup>e</sup> siècle que nous développerons en étudiant les travaux de Cayley sur les matrices.

décomposition polynomiale afin de résoudre le problème posé par l'occurrence de racines caractéristiques multiples. L'application de la méthode de Sylvester aux quadriques nécessite la prise en compte des seconds mineurs afin de caractériser les types d'intersections plus nombreux qui sont donnés en [encart 6](#). Le problème général du dénombrement des systèmes de mineurs extractibles d'un déterminant suscitera, en 1855, l'intérêt de Cayley pour la notion de matrice.

## 2. La théorie des matrices de Cayley.

### • Premiers travaux de Cayley sur les matrices de Sylvester.

Au début des années 1850, Cayley et Sylvester entretiennent une collaboration mathématique intense et travaillent notamment à jeter les bases de la théorie des invariants <sup>(19)</sup>. Dans un article publié en 1855 dans le *Journal de Crelle* et intitulé "Remarques sur la notation des fonctions algébriques", Cayley adopte pour la première fois la notion de matrice comme une "notation commode" pour représenter les systèmes linéaires et les formes quadratiques :

Un système de quantités rangées en forme de carré, mais d'ailleurs tout à fait indépendantes (je ne parle pas ici des matrices rectangulaires). Cette notation me paraît très commode pour la théorie des équations linéaires. ; j'écris par exemple

$$(\xi, \eta, \zeta, \dots) = \begin{pmatrix} \alpha, & \beta, & \gamma, & \dots \\ \alpha', & \beta', & \gamma', & \dots \\ \alpha'', & \beta'', & \gamma'', & \dots \\ \vdots & & & \end{pmatrix} (x, y, z, \dots)$$

pour représenter le système des équations

$$\begin{aligned} \xi &= \alpha x + \beta y + \gamma z, \dots, \\ \eta &= \alpha' x + \beta' y + \gamma' z, \dots, \\ \zeta &= \alpha'' x + \beta'' y + \gamma'' z, \dots, \\ &\vdots \end{aligned}$$

[Cayley, 1855b, 282].

Comme nous l'avons vu, les notions de matrices et mineurs avaient été introduites afin de caractériser les types d'intersection de deux coniques ou quadriques par la décomposition d'un déterminant en suites de mineurs. Généraliser la méthode à des situations comportant plus de trois ou quatre variables pose le problème du dénombrement des différentes décompositions possibles. Ce problème est à l'origine de l'intérêt de Cayley pour la notion de matrice et suscite la publication de trois articles successifs en 1855 <sup>(20)</sup>:

[...] ce déterminant sera une fonction de  $s$  du  $n^{\text{e}}$  degré qui généralement ne contiendra pas de facteurs multiples. On voit donc qu'un facteur simple du déterminant ne peut pas entrer comme facteur dans les premiers mineurs (c'est à dire dans tous les premiers mineurs) ; mais en supposant que le déterminant

<sup>19</sup> La proximité entre Arthur Cayley et James Joseph Sylvester est célèbre. La relation entre les deux hommes débute en 1846. Pour Sylvester qu n'a pas produit de résultat mathématique depuis son retour des Etats Unis en 1844, la rencontre avec Cayley semble vécue comme une renaissance : « Cayley, who, though younger than myself is my spiritual progenitor –who first opened my eyes and purged them of dross so that they could see and accept the higher mysteries of our common mathematical faith ».

<sup>20</sup> Ces articles, tous publiés en français dans le *Journal de Crelle*, sont les suivants : "Sur la Transformation d'une Fonction quadratique en elle-même par des Substitutions linéaires" [1855a], "Remarque sur le Notation des Fonctions algébriques" [1855b], "Recherches sur les matrices dont les termes sont des fonctions linéaires d'une seule indéterminée" [1855c]. Cayley donne la table explicite des différentes décompositions d'un déterminant en mineurs pour  $n < 6$  et associe la suite obtenue aux coefficients d'un développement en série entière de  $(1-x^j)^j$ . La méthode de Cayley permet notamment de corriger une erreur de décompte commise par Sylvester pour  $n=7$  et  $n=8$ .

ait des facteurs multiples, un facteur multiple peut entrer dans les premiers mineurs. Il importe de trouver le degré selon lequel un facteur multiple peut entrer comme facteur dans un des mineurs d'un ordre quelconque donné [...]. Cela se fait très facilement au moyen d'une propriété générale des déterminants : si les mineurs du  $(r+1)^e$  ordre contiennent le facteur  $(s-a)^\alpha$  (c'est à dire si tous les mineurs de cet ordre contiennent le facteur  $(s-a)^\alpha$ , mais non pas tous les facteurs  $(s-a)^{\alpha+1}$ ); et si de même les mineurs du  $r^{i\text{ème}}$  ordre contiennent le facteur  $(s-a)^\beta$  alors les mineurs du  $(r-1)^{i\text{ème}}$  ordre contiendront au moins le facteur  $(s-a)^{2\beta-\alpha}$ . Autrement dit : les mineurs du  $(r-1)^{i\text{ème}}$  ordre contiendront le facteur  $(s-a)^\gamma$  où  $\gamma \geq 2\beta - \alpha$ . [...] en formant la suite des indices des puissances selon lesquelles le facteur  $(s-a)$  entre dans les mineurs premiers, seconds, &c. (il va sans dire que cette suite sera une suite décroissante), les différences secondes seront positives [...]. Je représente par  $\alpha, \beta, \gamma, \dots$  la suite dont il s'agit ; je suppose, pour fixer les idées, que  $\delta$  soit le dernier terme qui ne s'évanouisse pas et j'écris

$$\begin{array}{ccccccc} \alpha & \beta & \gamma & \delta & 0 & 0 & \dots \\ \alpha-\beta & \beta-\gamma & \gamma-\delta & \delta & 0 & \dots & \\ \alpha-2\beta+\gamma & \beta-2\gamma+\delta & \gamma-2\delta & \delta & 0 & \dots & \end{array}$$

Ici, quel que soit le nombre des termes, tous les nombres de la troisième ligne seront positifs, et en représentant ces nombres par  $f, f', f'', \dots$ , on obtient :

$$\begin{array}{l} \alpha = f + 2f' + 3f'' + 4f''' + \dots ; \\ \beta = f' + 2f'' + 3f''' + \dots ; \\ \gamma = f'' + 2f''' + \dots ; \\ \delta = f''' + \dots ; \end{array}$$

Il y a ici à considérer que le nombre  $\alpha$ , indice de la puissance selon laquelle le facteur  $(s-a)$  entre dans le déterminant, est donné; il sera donc permis de prendre pour  $f, f', f'', \dots$  des valeurs entières et positives quelconques (zéro y compris) qui satisfont à la première équation [...] on forme de cette manière une table des particularités que peut présenter un facteur multiple  $(s-a)^\alpha$  du déterminant [...]. [Cayley, 1855c, 313].

### • Le "théorème remarquable" de Cayley.

La publication en 1858 du mémoire intitulé « A Memoir on the Theory of Matrices », manifeste une évolution du point de vue de Cayley sur la notion de matrice. La matrice n'est plus, en 1858, une simple notation commode permettant de distinguer un objet comme un système linéaire ou une forme quadratique de son déterminant. Elle fait désormais l'objet d'une « théorie » s'articulant autour de l'énoncé d'un "théorème remarquable"<sup>(21)</sup> :

I obtain the remarkable theorem that any matrix whatever satisfies an algebraical equation of its own order, the coefficient of the highest power being unity, and those of the other powers functions of the terms of the matrix, the last coefficient being in fact the determinant; the rule for the formation of this equation may be stated in the following condensed form, which will be intelligible after a perusal of the memoir, viz. The determinant formed out of the matrix diminished by the matrix considered as a single quantity involving the matrix unity, will be equal to zero. [Cayley, 1858, 17].

Traduction. F.B.

J'obtiens le théorème remarquable selon lequel toute matrice satisfait une équation algébrique de son propre ordre, le coefficient de la plus grande puissance étant égal à l'unité, ceux des autres puissances étant fonctions des termes de la matrice et le dernier coefficient étant égal au déterminant ; la loi pour la formation de ces équations peut être énoncé sous la forme condensée suivante, qui deviendra intelligible après étude du mémoire : le déterminant formé par une matrice soustraite à la matrice considéré comme une quantité simple associant la matrice unité, est égal à zéro.

Le mémoire de Cayley sera beaucoup cité et commenté par les mathématiciens travaillant, à la fin du XIX<sup>e</sup> siècle, sur la théorie des systèmes hypercomplexes. A partir de 1890, il sera perçu comme un des premiers travaux à l'origine de la théorie des algèbres associatives

<sup>21</sup> Remarquer le terme essentiel de « single quantity involving the matrix unity » utilisé par Cayley et sur lequel nous aurons l'occasion de revenir plus loin.

Nous reviendrons dans la conclusion de l'article sur la dénomination e « théorème de Cayley-Hamilton » donnée entre 1890 et 1900.

(encart 7). Pour cette raison, l'historiographie a porté un regard insistant sur les « lois » d'opérations sur les matrices établies par Cayley (encart 12), ces lois ayant souvent été présentées comme une des origines de *l'algèbre associative des matrices* (<sup>22</sup>). Le mémoire de 1858 n'est pourtant pas motivé par une volonté d'abstraction visant à généraliser les opérations de l'arithmétique à des objets *hypercomplexes* de plusieurs dimensions. Le « théorème remarquable » qui fonde la théorie de Cayley est en effet énoncé pour résoudre *un problème mathématique précis*, à savoir *l'expression des fonctions rationnelles des fonctions homographiques*. Dès 1855, Cayley avait remarqué la possibilité d'employer la notation matricielle pour représenter les « fonctions linéaires » intervenant dans la « théorie des figures homographes » :

Mais tout cela [la notation des matrices] s'applique à une autre théorie géométrique, savoir à la théorie des figures homographes. Pour fixer les idées, je ne considère que les figures dans le plan. En supposant que  $x, y, z$  soient les coordonnées d'un point, et en prenant pour  $(X, Y, Z)$  des fonctions linéaires de  $(x, y, z)$  on aura  $(X, Y, Z)$  comme coordonnées d'un point homographe au point  $(x, y, z)$ . En cherchant les points qui sont homographes chacun à soi même, on est conduit aux équations

$$X-sx=0, Y-sy=0, Z-sz=0.$$

Les quantités à gauche  $X-sx, Y-sy, Z-sz$  sont des fonctions linéaires de  $x, y, z$  ayant pour coefficients des fonctions linéaires de  $s$ . On a ainsi une matrice dont les termes sont des fonctions linéaires de  $s$  ; la théorie entière se rattache aux propriétés de cette matrice. [...] Je reviendrai à cette théorie à une autre occasion. [Cayley, 1855b, 282].

Si l'on compose une fonction homographique par elle même, on obtient ce que l'on peut considérer comme une fonction homographique *au carré*. Babbage, un prédécesseur de Cayley, s'était déjà préoccupé du problème de l'expression des racines carrées des fonctions homographiques et le véritable objet de la théorie des matrices de 1858 est d'énoncer un théorème supportant une *méthode de calcul des fonctions rationnelles des homographies* :

[...] it is nevertheless possible to form the powers (positive or negative, integral or fractional), of a matrix, and hence to arrive at the notion of a rational and integral function, or generally of any algebraical function, of a matrix. I obtain the remarkable theorem that any matrix whatever satisfies an algebraical equation of its own order [...]. The theorem shows that every rational and integral function (or indeed every rational function) of a matrix may be considered as a rational and integral function, the degree of which is at most equal to that of the matrix, less unity: it even shows that in a sense, the same is true with respect to any algebraical function whatever of a matrix. [Cayley, 1858, 17].

Traduction F.B.

On peut former les puissances (positives ou négatives, entières ou fractionnaires), d'une matrice et obtenir ainsi la notion de fonction entière, rationnelle, ou plus généralement de fonction algébrique d'une matrice. J'obtiens le théorème remarquable selon lequel toute matrice satisfait une équation algébrique de son propre ordre [...]. Le théorème montre que toute fonction rationnelle d'une matrice peut être considérée comme une fonction rationnelle de degré au plus égal à celui de la matrice moins une unité : en un sens, la même chose est vraie pour une fonction algébrique quelconque d'une matrice.

- **L'héritage de l'« école algébrique anglaise ».**

Plutôt que de rechercher dans la théorie des matrices de Cayley des prémisses de structures algébriques comme les algèbres associatives, le mémoire de 1858 gagne à être mis en perspective des travaux de la génération de mathématiciens qui précède Cayley et Sylvester et que l'on désigne souvent comme l'« *école algébrique anglaise* » de la première moitié du XIX<sup>e</sup>

---

<sup>22</sup> C'est plus précisément la définition de l'addition des matrices qui a souvent été considérée comme un pas supplémentaire vers l'« abstraction » d'une théorie des algèbres associatives (la multiplication matricielle s'interprète quant à elle comme la composition des transformations linéaires). Voir les passages consacrés à la notion de matrice par B.L. Van der Waerden [1977], T. Hawkins, [1975, 1977], J. Dieudonné [1978], K. Parshall [1988], I. Grattan-Guinness [1994] ou L. Corry [1996].

siècle <sup>(23)</sup>. L'objet du mémoire de Cayley, la généralisation de fonctions rationnelles à des expressions symboliques, était en effet une des grandes préoccupations de l'école algébrique anglaise. La traduction du *Traité élémentaire de calcul différentiel et intégral* de Lacroix [1802] par Babbage est souvent présentée comme exemplaire des efforts du réseau qui se développe autour d'un groupe d'étudiants de Cambridge pour introduire en Angleterre les méthodes de l'*algèbre analytique* de Lagrange. Les méthodes de Lagrange et, plus généralement celles du *calcul différentiel* développé dans la tradition de Leibniz, étaient alors critiquées en Angleterre pour le caractère *mécanique* de leurs procédures et les *paradoxes* sur les quantités impossibles qu'elles impliquaient. A ces critiques, les membres du "network" de Cambridge opposaient une *philosophie de l'algèbre* qui se caractérisait par l'importance donnée aux *opérations* sur les objets. L'"algèbre symbolique" de Peacock de 1833 établissait une *distinction* entre *signification* et *symboles* et impliquait une rupture avec "le réalisme mathématique du XVIII<sup>e</sup> siècle selon lequel à tout objet mathématique correspond dans la réalité un élément essentiel qui en constitue la légitimité" [Durand-Richard, 1990, 131]. L'historienne Marie-José Durand-Richard a éclairé la spécificité conceptuelle de l'approche symbolique anglaise par le contexte historique de l'Angleterre de la révolution industrielle et de la philosophie de Locke [Durand-Richard, 1996].

Sur le modèle du mémoire de Morgan de 1841, *On the foundations of Algebra*, Cayley explore, dans sa théorie des matrices, des propriétés de *procédés symboliques* et définit les *lois opératoires* qui donnent à la matrice-mère des mineurs de Sylvester un comportement similaire à celui des « quantités algébriques ordinaires » ([encart 12](#)).

It will be seen that matrices (attending to those of the same order) comport themselves as single quantities; they may be added, multiplied or compounded together, etc. : the law of addition of matrices is precisely similar to that for the addition of ordinary algebraical quantities ; as regards to their multiplication (or composition) there is the peculiarity that matrices are not in general convertible. [Cayley, 1858, 17].

Traduction. F.B.

Nous verrons que les matrices (toutes prises du même ordre) se comportent comme des quantités simples : elles peuvent être ajoutées, multipliées ou composées les unes avec les autres : la loi d'addition des matrices est précisément similaire à celle de l'addition des quantités algébriques ordinaires, en ce qui concerne la multiplication (ou la composition) il y a une spécificité car les matrices ne sont pas en général échangeables.

L'importance que prennent, dans le mémoire de Cayley, les problèmes de calculs de puissances et racines de matrices renvoie à une préoccupation *traditionnelle* de l'école anglaise depuis les travaux de Herschell, en 1813, sur la *notation des opérateurs différentiels* et la propriété  $f^n(f(x)) = f^{n+1}(x)$ . La partie du mémoire de Cayley consacrée aux "matrices périodiques", et plus généralement aux équations de matrices, n'a pas été commentée par l'historiographie qui s'est concentrée sur la définition des opérations sur les matrices. *Il s'agit pourtant là de l'objet même du mémoire de Cayley : montrer que toute fonction rationnelle d'une matrice, en particulier  $\sqrt{M}$ , peut s'exprimer comme une "fonction entière" de degré inférieur à celui de la matrice elle-même.* C'est pour résoudre ce problème que Cayley définit les opérations sur les matrices qui s'avèrent nécessaires pour exprimer  $\sqrt{M}$  par des fonctions entières.

- **La notion de « single quantity » et les pratiques spécifiques au calcul des matrices de Cayley.**

---

<sup>23</sup> L. Novy [1968, 211-222] a le premier qualifié d'école algébrique anglaise le courant de pensée dont les acteurs principaux sont C. Babbage, G. Peacock, D. Gregory, A. de Morgan, W.R. Hamilton et G. Boole.



En 1858, Cayley ne se préoccupe pas du problème des racines caractéristiques multiples qui motivait Sylvester, mais cherche à *exprimer des fonctions rationnelles d'homographies*. A cette fin, il élabore des *pratiques* qui reposent sur la nature duale de la notion de matrice à la fois *système de nombres* et *quantité simple*. Ces pratiques qui se manifestent notamment dans la démonstration du théorème remarquable. Partant d'une matrice

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

Cayley forme le déterminant

$$\begin{vmatrix} a-M & b \\ c & d-M \end{vmatrix}.$$

Le calcul effectif des puissances de  $M$  permet de montrer que le développement de ce déterminant,

$$M^2 - (a+d)M^1 + (ad-bc)M^0,$$

est nul, ce qui démontre le théorème remarquable pour  $n=2$ . Mais que signifie ce déterminant où se trouvent ajoutés un symbole-nombre  $a$  et un symbole-matrice  $M$  ? Pour Cayley, il s'agit du déterminant dont la matrice est

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot M \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Dans cette écriture la notation  $M$  désigne à la fois une *quantité multiple* et le produit d'une *quantité*  $M$  par la matrice unité. Cayley nomme "single quantity" la matrice du membre de droite résultant du produit d'une quantité  $m$  par la matrice unité :

$$m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}$$

Le symbole  $\tilde{M}$  permet de distinguer la "quantité" de la matrice "single quantity" correspondante :

[...] let the matrix  $M$ , considered as a single quantity, be represented by  $\tilde{M}$ , then writing 1 to denote the matrix unity,  $\tilde{M} \cdot 1$  will represent the matrix  $M$ , considered as a single quantity involving the matrix unity. [Cayley, 1858, 30].

Traduction F.B.

[...] représentons par  $\tilde{M}$  la matrice  $M$  considérée comme "single quantity", alors, dénotant par 1 la matrice unité,  $\tilde{M} \cdot 1$  représentera la matrice  $M$ , considérée comme une « single quantity » associant la matrice unité.

La notion de «single quantity» associée à la matrice unité permet à Cayley d'énoncer le théorème remarquable sous la forme condensée suivante :

[...] le déterminant formé par une matrice soustraite à la matrice considéré comme une quantité simple associant la matrice unité, est égal à zéro.

Comme nous allons le voir en suivant la méthode développée par Cayley pour l'étude des matrices périodiques vérifiant  $M^2=I$ , la notion de « single quantity » supporte une *pratique de factorisation d'équations polynomiales dont les coefficients sont interprétés tantôt comme des nombres, tantôt comme des « single quantity », c'est-à-dire des matrices réduites à une diagonales répétant l'occurrence d'un même nombre*.

[...]; it will be seen presently that the equation  $M^2=I$  admits of other solutions besides  $M=\pm I$ . The example shows how the values of the fractional powers of a matrix are to be investigated. [Cayley, 1858, 485].

L'équation  $M^2=I$  admet d'autres solutions que  $\pm I$ . C'est pour l'étude de cette propriété particulière des équations de matrices que Cayley développe une méthode basée sur l'emploi du théorème remarquable:

But suppose it is required to find a matrix of the order 3,

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

which shall be periodic of the second order. Writing for shortness

$$\begin{vmatrix} a-M & b & c \\ d & e-M & f \\ g & h & i-M \end{vmatrix} = (-M^3-AM^2+BM-C),$$

the matrix here satisfies

$$M^3-AM^2+BM-C=0$$

[Cayley, 1858, 487].

La méthode de Cayley consiste à factoriser l'équation  $M^3-AM^2+BM-C=0$  par  $M^2-I$ . Si  $I+B=0$  et  $A+C=0$ , on obtient :

$$(M^2-I)(M+C)=0$$

Comme le fait remarquer Cayley, on aimerait déduire de cette dernière équation la conclusion  $M^2-I=0$  et  $M^2=I$ . Or, et c'est tout l'objet de la dernière partie du mémoire de Cayley, on ne peut déduire de l'équation en matrices  $(M^2-I)(M+C)=0$ , l'équation  $M^2-I=0$  que dans le cas où la matrice  $M+C$  a un déterminant non nul:

There is an apparent difficulty connected with the equation satisfied by a matrix, which it is proper to explain. Suppose, as before,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

so that  $M$  satisfies the equation

$$\begin{vmatrix} a-M & b \\ c & d-M \end{vmatrix} = 0$$

or

$$M^2-(a+d)M+ad-bc=0$$

and let  $X, X'$  be the single quantities, roots of the equation

$$\begin{vmatrix} a-X & b \\ c & d-X \end{vmatrix} = 0$$

or

$$X^2-(a+d)X+ad-bc=0.$$

The equation satisfied by the matrix may be written

$$(M-X')(M-X)=0,$$

in which  $X, X'$  are to be considered as respectively involving the matrix unity, and it would at first sight seem that we ought to have one of the simple factors equal to zero; this is obviously not the case, for such equation would signify that the perfectly indeterminate matrix  $M$  was equal to a single quantity, considered as involving the matrix unity. The explanation is that each of the simple factors is an indeterminate matrix; in fact [...] the determinant of this matrix is equal to zero. the product of the two factors is thus equal to zero without either of the factors being equal to zero.

[1858, 485].



Nous voyons la notion de matrice adopter une *nouvelle identité* dans le mémoire de Cayley en 1858. Une matrice ne se caractérise plus comme la mère des mineurs d'un déterminant mais par les *lois d'un calcul symbolique* et l'énoncé du *théorème remarquable*. Comme nous l'avons vu en mentionnant l'héritage de l'*école algébrique anglaise*, l'*identité* des matrices de Cayley est indissociable d'un *contexte culturel* spécifique. Cette *identité* est également associée à des *pratiques spécifiques de factorisations de polynômes de matrices* qui reposent sur une conception des matrices comme des quantités tout à la fois *simples et multiples*. Nous reviendrons sur la *spécificité* de cette pratique en posant la question, dans la troisième partie de cet article, de son *héritage* dans les mathématiques de la fin du XIX<sup>e</sup> siècle.

### III. Evolutions de la notion de matrice entre 1850 et 1890.

Entre la publication du mémoire de 1858 et les années 1890, le terme « matrice » disparaît presque entièrement des textes mathématiques. Dans cette même période la notion de « mineur », pourtant introduite simultanément par Sylvester, est adoptée par de nombreux mathématiciens parmi lesquels Hermite, Jordan, Darboux ou encore Poincaré <sup>(24)</sup>. Des graphes des réseaux de textes comme celui porté en [encart 4](#) montrent le mémoire de Cayley s'éclipser une trentaine d'années avant de réapparaître dans un enchevêtrement de références bibliographiques dans les années 1890. Les auteurs qui emploient la notion de matrice entre 1890 et 1900 citent systématiquement le mémoire publié par Cayley en 1858 comme un texte fondateur célébré pour sa définition d'une algèbre associative des matrices. A ce rôle fondateur attribué au mémoire de 1858, il faut opposer l'absence quasi complète de la notion de matrice durant la trentaine d'années qui suit la publication du mémoire <sup>(25)</sup>. Comment interpréter cette éclipse dans l'héritage de la théorie des matrices de Cayley ? Comment le texte de Cayley acquiert-il une postérité si forte après des décennies d'indifférence ? Ces deux questions se mêlent en une seule : la notion de matrice de 1890 est-elle la même que celle de Cayley de 1858 ? *Il faut en fait, entre la définition de 1858 et l'héritage de 1890, reconnaître une double origine historique de la notion de matrice.* Aborder cette double origine revient à considérer l'évolution de la notion de matrice dans le temps.

#### 1. Héritages de la théorie des matrices de Cayley dans les travaux de Sylvester de la période 1882-1885

- **Quand Sylvester (1882) lit Cayley (1858) : un problème et sa résolution, le calcul de  $\sqrt{M}$ .**

Trente années séparent la première définition de la matrice comme *mère des mineurs* et sa réapparition dans l'œuvre de Sylvester à l'occasion d'une note aux *Comptes Rendus de l'Académie des Sciences* datée de février 1882. Il s'agit de résoudre un problème déjà présenté dans une note publiée dans les *Comptes Rendus* le 9 janvier de la même année et intitulée "Sur les puissances et les racines des substitutions linéaires". Sylvester généralise aux fonctions homographiques d'un nombre quelconque de variables un problème classique, déterminer une fonction homographique  $\varphi$  de "périodicité donnée" <sup>(26)</sup>, c'est-à-dire déterminer

$$\varphi(x) = \frac{ax + \alpha}{bx + \beta}$$

telle que

$$\varphi^\mu(x) = x.$$

<sup>24</sup> Hermite semble à l'origine de l'adoption du terme « mineur » en France. Dans les années 1850, Hermite suit de près les travaux de Cayley et Sylvester sur les invariants et adopte la notion de mineur très rapidement. En Allemagne, les travaux des Anglais suscitent l'intérêt de Riemann qui utilise la notion de matrice de Sylvester en 1857 pour représenter des systèmes d'équations différentielles linéaires et en extraire des mineurs.

<sup>25</sup> Même en Angleterre, la notion de matrice est peu utilisée avant les années 1880-1890, H.J.S. Smith est l'un des seuls mathématiciens à faire référence à la théorie des matrices de Cayley (entre 1861 et 1863).

<sup>26</sup> Ce problème, comme nous l'avons vu, était traité par Cayley en 1858. L'intérêt de Sylvester pour ce problème provient de ses récents travaux, parus dans le *Journal de Mathématiques* [1881a], sur l'inégalité de Tchebycheff [1852] donnant des bornes aux nombres de nombres premiers  $\pi(x)$  inférieurs à un nombre donné  $x$  :

$$A_1 < \pi(x)/x/\log(x) < A_2 \\ \text{où } 0,922 < A_1 < 1 \text{ et } 1 < A_2 < 1,105.$$

Les calculs de Sylvester nécessitent des systèmes d'équations linéaires dont les solutions reviennent à rechercher les racines d'une fonction homographique. Voir à ce sujet [Parshall, 1985].

La méthode de Sylvester consiste à écrire *les fonctions puissances et racines de la substitution linéaire associée à  $\varphi$  comme une fonction numérique des "racines lambdaïques" de l'équation  $\det(\varphi - \lambda I) = 0$*  :

Soit un déterminant quelconque donné, et ajoutons le terme  $-\lambda$  à chaque terme diagonal ; on obtient ainsi une fonction de  $\lambda$  ; je nomme les racines de cette fonction racines *lambdaïques* du déterminant donné [...].  $i$  étant une quantité commensurable quelconque, les  $i^{\text{èmes}}$  puissances des racines lambdaïques d'un déterminant de substitution sont identiques avec les racines lambdaïques de  $i^{\text{ème}}$  puissance du déterminant.  
[Sylvester, 1882a, 56]

La note de janvier, qui ne fait aucune référence à la notion de matrice, est suivie par la parution, au mois de février, d'une seconde note intitulée "Sur les racines des matrices unitaires" et dans laquelle le problème est reformulé de la manière suivante :

Extraire la racine  $\mu^{i^{\text{ème}}}$ , ou plus généralement trouver la puissance  $i^{\text{ème}}$  d'une matrice donnée. [1882a, 57].

La notion de matrice apparaît en février 1882 afin de corriger une erreur commise par Sylvester dans sa précédente note. La formule donnée pour exprimer les fonctions de substitutions comme fonctions numériques des "racines lambdaïques" est fautive dans le cas où les "racines lambdaïques" *ne sont pas toutes distinctes* <sup>(27)</sup> :

$$K_i = \sum \frac{K_{n-1} - S_1 \cdot K_{n-2} + S_2 \cdot K_{n-3} - \dots \pm S_{n-2} \cdot K_1 \mp S_{n-1} \cdot K_0}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \dots (\lambda_1 - \lambda_n)} \lambda_i^i,$$

Dans la note de janvier, Sylvester pensait pouvoir traiter l'occurrence de *racines multiples* par la simple "introduction de différences infinitésimales". La formule obtenue est cependant en contradiction avec un résultat donné par Cayley en 1858 pour la résolution de l'équation  $M^3 = I$  [Cayley, 1858, 487-488]. Les difficultés rencontrées par Sylvester présentent des similitudes avec celles résolues en 1851 par l'introduction des notions de *matrices* et de *mineurs* ; dans les deux cas, la difficulté provient de l'occurrence de *racines multiples* dans l'équation  $\det(\varphi - \lambda I) = 0$ . Comme nous l'avons vu dans la partie II, Sylvester avait montré en 1851 qu'il était nécessaire d'aller au-delà de l'examen du déterminant et de considérer les mineurs de la matrice associée.

<sup>27</sup> Au sujet de la formule de Sylvester, voir l'exposé de Cartan et Study dans *l'Encyclopédie des sciences mathématiques* [1908, 438] sur l'expression des "fonctions analytiques" dont les variables sont hypercomplexes (par exemple les exponentielles de matrices). Voir aussi les traités de théorie des matrices des années 1930 et notamment la formulation donnée au problème par Wedderburn. Il s'agit alors de déterminer un polynôme de matrices et, dans le cas où la matrice  $x$  n'a que des valeurs propres simples, des polynômes numériques suffisent à caractériser le polynôme  $g(x)$ ; l'occurrence de racines multiples nécessite l'emploi de la forme de Jordan. Wedderburn expose d'abord le cas particulier d'occurrence d'une seule racine multiple [Wedderburn, 1931, 27]:

Si l'on suppose que la matrice  $x$  a une seule racine caractéristique, alors son equation minimale est :

$$\varphi(\lambda) = (\lambda - \lambda_1)^v,$$

Posons

$$\eta_i^j = \eta_i = (x - \lambda_1)^j = (x - \lambda_1)_{\eta_{i-1}}$$

alors

$$\begin{aligned} \eta_1^v &= 0, \quad x\eta_{v-1} = \lambda_1 \eta_{v-1}, \quad x\eta_i = \lambda_1 \eta_i + \eta_{i+1}, \\ [\dots] g(x) &= g(\lambda_1 + \eta_1) = g(\lambda_1) + g'(\lambda_1)\eta_1 + \dots + \frac{g^{(v-1)}(\lambda_1)}{(v-1)!} \eta_1^{v-1}. \end{aligned}$$

Dans le cas général où la fonction caractéristique est

$$\varphi(\lambda) = \prod_{i=1}^r (\lambda - \lambda_i)^{v_i}$$

( $\sum v_i = v$ ,  $r > 1$ ), alors :

$$g(x) = \sum_{i=1}^r [g(\lambda_i)\varphi_i + g'(\lambda_i)\eta_i + \dots + \frac{g^{(v_i-1)}(\lambda_i)}{(v_i-1)!} \eta_i^{v_i-1}(\lambda)]$$

On peut demander quelle est la forme d'une autre matrice  $M$  du même ordre  $n$ , telle que la  $i^{\text{ème}}$  puissance de  $M$  soit une matrice unitaire. [...] Je vais à présent donner toutes les solutions dont la question est susceptible. Soient  $v_1, v_2, v_3, \dots, v_k$  des nombres entiers et positifs quelconques dont la somme est  $n$ , et  $\rho_1, \rho_2, \dots, \rho_k$   $k$  quelconques des  $i^{\text{ème}}$  racines de l'unité. Soit  $M_\lambda$  la matrice affectée de l'indice  $\lambda$ , c'est-à-dire modifiée par l'addition de  $-\lambda$  à chacun des  $n$  termes de la diagonale. Considérons les systèmes de matrices mineurs de  $M$ , de l'ordre  $n-v_1+1, n-v_2+1, \dots, n-v_k+1$  respectivement ; et prenons  $M$  tel que  $\lambda-\rho_1, \lambda-\rho_2, \dots, \lambda-\rho_k$  soient facteurs de chaque mineur du premier, du second, ..., du  $k^{\text{e}}$  de ces systèmes respectivement : alors  $M$  sera une racine  $i^{\text{ème}}$  de la matrice unitaire de l'ordre  $n$ . [1882b, 396].

En cas d'occurrence de racines multiples, la notion de matrice comme génératrice des mineurs permet de donner toutes les solutions et d'accorder le résultat de Sylvester avec le cas particulier traité par Cayley en 1858 :

Ainsi, pourvu que  $i$  soit égal ou supérieur à  $n$ , il y aura autant de genres de racines  $i^{\text{èmes}}$  de cette matrice qu'il a de partition indéfinies de  $n$ .

[...] En effet, la matrice trouvée par M. Cayley, dans son Mémoire sur les matrices (*Philosophical Transactions*, 1858),

$$\begin{array}{ccc} \frac{\alpha}{\alpha + \beta + \gamma}, & \frac{-(\beta + \gamma) \nu}{\alpha + \beta + \gamma \mu}, & \frac{-(\beta + \gamma) \nu}{\alpha + \beta + \gamma \mu}, \\ \frac{-(\gamma + \alpha) \mu}{\alpha + \beta + \gamma \nu}, & \frac{\beta}{\alpha + \beta + \gamma}, & \frac{-(\gamma + \alpha) \lambda}{\alpha + \beta + \gamma \mu}, \\ \frac{(\alpha + \beta) \mu}{\alpha + \beta + \gamma \nu}, & \frac{-(\alpha + \beta) \nu}{\alpha + \beta + \gamma \lambda}, & \frac{\gamma}{\alpha + \beta + \gamma}, \end{array}$$

sera la matrice  $M$ , telle que chaque mineur de  $M_\rho$  contiendra  $(\rho-1)$  ; de même chaque mineur de  $(-M)_\rho$  contiendra  $\rho+1$  ; on remarquera que 1 et -1 sont les racines carrées de l'unité, et l'on vérifiera aisément que  $M^2$  ou, ce qui revient au même,  $\Phi(-M)^2$  ont tous les deux la forme

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

[Sylvester, 1882b, 397].

La réapparition de la notion de matrice en 1882 est donc avant tout le résultat d'une relecture par Sylvester de ses propres travaux de 1851. La référence à Cayley a le même statut que les références à Babbage et Serret, elle vient citer le travail d'un prédécesseur sur le problème du *calcul des puissances d'une substitution*. Nous avons vu en effet que la *motivation initiale* de la théorie des matrices de Cayley visait une méthode pour déterminer des *fonctions rationnelles de fonctions homographiques*, le cas de la racine carrée  $\sqrt{M}$  correspondant à l'exemple le plus classique, associé au nom de Babbage.

- **La seconde lecture de Cayley par Sylvester : les matrices comme quantités complexes.**

Les difficultés de la résolution des équations en matrices, amènent Sylvester à relire le mémoire de Cayley. La lecture par Sylvester du mémoire de 1858 et sa conception de la notion de matrice évoluent radicalement entre 1882 et 1883.

Je terminerai en ajoutant que j'ai établi une théorie fonctionnelle générale des matrices, et que je ne regarde plus celles-ci comme des schemata d'éléments, mais comme des communautés ou, si l'on veut, comme des quantités complexes. [Sylvester, 1882b, 398].

Alors qu'en février 1882, Sylvester conservait sa définition d'une matrice comme mère des mineurs d'un déterminant, c'est une "théorie des matrices" participant d'une "algèbre multiple" qui est présentée dans un mémoire adressé en 1883 au *Philosophical Magazine*.

Pour Sylvester la "nouvelle méthode" élaborée en 1882 pour extraire la racine  $\mu^e$  d'une substitution donnée, permet d'énoncer un "théorème général" participant d'une *théorie "universelle" car provoquant le "rapprochement inattendu" de questions diverses comme les théories des invariants, des systèmes d'équations différentielles, des fonctions homographiques et des quaternions* :

Of the many unexpected results which I have obtained by my new method, not the least striking is the *rapprochement* which it establishes between the theory of Matrices and that of Invariants. The theory of invariance relative to associated Matrices includes and transcends that relative to algebraical functions.

[...] Babbage's famous investigation of the form of the homographic function of  $\frac{px+q}{rx+s}$  of  $x$ , which has a periodicity of any given degree  $q$ , is in fact (surprising as such a statement would have appeared to Babbage and Hamilton) one and the same thing as to find the  $q$ th root of unity under the form of a quaternion ! [Sylvester 1883a, 111-114].

Traduction. F.B.

Le rapprochement entre la théorie des Matrices et celle des Invariants n'est pas le moins frappant des résultats inattendus que j'ai obtenu à l'aide de ma nouvelle méthode. La théorie de l'invariance relative aux Matrices associées inclus et transcende celle relative aux fonctions algébriques. Les fameuses recherches de Babbage sur la forme de la fonction homographique  $\frac{px+q}{rx+s}$  de  $x$ , de périodicité un degré  $q$  donné, reviennent en fait, aussi surprenant que cela pourrait paraître à Babbage ou Hamilton, à la même chose que de trouver la racine  $q^e$  de l'unité de la forme d'un quaternion !

La "nouvelle méthode" est basée sur l'introduction d'invariants, les "racines lambdaïques" de 1882 rebaptisées "racines latentes", qui permettent de caractériser des *fonctions de matrices* par des *fonctions numériques* :

It will be convenient to introduce here a notion (which play a conspicuous part in my new theory of multiple algebra), namely that of the *latent roots* of a matrix – latent in a somewhat similar sense as vapour may be said to be latent in water or smoke in a tobacco leaf. If from each term in the diagonal of a given matrix,  $\lambda$  be subtracted, the determinant to the matrix so modified will be a rational integer function of  $\lambda$ ; the roots of that function are the latent roots of the matrix, and there results the important theorem that the latent roots of any function of a matrix are respectively the same functions of the latent roots of the matrix itself: for example, the latent roots of the square of a matrix are the square of its latent roots [1883a, 110].

Traduction. F.B.

Il sera commode d'introduire la notion de *racine latente* d'une matrice – latente en un sens similaire à la vapeur qui peut être dire latente dans l'eau ou à la fumée latente dans une feuille de tabac. Si l'on soustrait  $\lambda$  de chaque terme de la diagonale d'une matrice, le déterminant de la matrice obtenue sera une fonction rationnelle entière de  $\lambda$  ; les racines de cette fonctions sont les racines latentes de la matrices et on obtient le théorème important selon lequel les racines latentes d'une fonction quelconque de matrices sont respectivement les mêmes fonctions des racines latentes de la matrice elle même : par exemple, les racines latentes du carré d'une matrice sont les carrés des racines latentes de la matrice.

Basée sur les résultats publiés dans les Comptes Rendus en 1882, la "seconde loi du mouvement algébrique" de Sylvester exprime toute fonction de matrices,  $\phi(m)$ , comme une fonction des racines latentes de  $m$ ,  $\phi(\lambda)$  [Sylvester 1883b, 114] :

$$\phi m = \sum \frac{(m - \lambda_2)(m - \lambda_3) \dots (m - \lambda_i)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \dots (\lambda_1 - \lambda_i)} \phi \lambda.$$

Sylvester considère désormais une matrice comme "une quantité complexe" au sens de la "single quantity" de Cayley qui permet, comme nous l'avons vu dans la deuxième partie, de concevoir les symboles  $\lambda_i$  tout à la fois comme des nombres (les racines latentes de  $m$ ) et des matrices (diagonales). L'énoncé de la "seconde loi du mouvement algébrique" nécessite en effet l'introduction du calcul symbolique des matrices, et notamment de la règle d'addition:

This theorem of course presupposes the rule first stated by Prof. Cayley (*Phil. Trans.* 1857) for the addition of matrices. [1883, 111].

L'application de la loi du mouvement algébrique au calcul des racines de quaternions appuie la revendication d'universalité de la nouvelle théorie, l'"algèbre multiple", à laquelle sera consacrée la publication d'une nouvelle série de notes aux Comptes Rendus en 1884 :

Qu'il me soit permis, avant de conclure, d'ajouter encore une petite réflexion sur l'importance de la question traitée ici. Elle constitue, pour ainsi dire, un canal qui, comme celui de Panama, sert à unir deux grands océans, celui de la théorie des invariants et celui des quantités complexes ou multiples : dans l'une de ces théories, en effet, on considère l'action des substitutions sur elles-mêmes, et dans l'autre, leur action sur les formes; de plus, on voit que la théorie *analytique* des quaternions, étant un cas particulier de celle des matrices, cesse d'exister comme une science indépendante; ainsi, de trois branches d'analyse autrefois regardées comme étant indépendantes, en voilà une abolie ou absorbée, et les deux autres réunies en une seule de substitution algébrique. [Sylvester, 1884].

- **De nouvelles bases pour la théorie des quaternions.**

En cherchant à "absorber" la théorie des quaternions dans celle des matrices, Sylvester est amené à donner un nouveau rôle à la représentation matricielle. Il s'avère en effet nécessaire de représenter par des matrices les "quaternions fondamentaux" et c'est la *forme des matrices* elle-même qui permet de caractériser les éléments de base d'un système de "grandeurs généralisées". Pour déterminer la *forme matricielle* des éléments de base, il suffit de résoudre des *équations* données par les "lois" définissant un système de grandeurs généralisées.

ON sait qu'on peut tout à fait (et très avantageusement) changer la base de la théorie des quaternions en considérant les trois symboles  $i, j, k$  de Hamilton comme des matrices binaires.

Si  $h, j$  sont des matrices binaires qui satisfont à l'équation  $hj = -jh$ , on démontre facilement que, en écartant le cas où  $hj = jh = 0$ ,  $A^2$  et  $k^2$  seront de la forme

$$\begin{pmatrix} c & 0 & \gamma & 0 \\ 0 & c' & 0 & \gamma \end{pmatrix}$$

c'est-à-dire  $cu, \gamma u$ , où  $u$  est l'unité binaire

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

On peut ajouter, si l'on veut, les deux conditions  $c^2 = 1, \gamma^2 = -1$ ; alors, en supprimant, pour plus de brièveté, le  $u$ , qui jouit de propriétés tout à fait analogues à celles de l'unité ordinaire, on obtient facilement les équations connues

$$h^2 = \bar{1}, \quad j^2 = \bar{1}, \quad k^2 = \bar{1}, \\ hj = -jh = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

[...] Une solution, parmi les plus simples des équations  $ij = -ji$ ;  $i^2 = -1, j^2 = -1$ , est la suivante :

$$i = \begin{pmatrix} \theta & 0 \\ 0 & -\theta \end{pmatrix}, \quad j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

et conséquemment

$$k = ij = \begin{pmatrix} 0 & -\theta \\ -\theta & 0 \end{pmatrix},$$

où  $\theta = \sqrt{-1}$ . [Sylvester, 1883c, 1337] :

La série de notes publiées aux *Comptes Rendus* entre 1884 et 1885 a pour objet de définir des "nonions" dont les éléments de base sont définis par des matrices d'ordre 3 :

On peut construire d'une manière tout à fait analogue un système de nonions en considérant l'équation  $m = \rho n$ , où  $m, n$  sont des matrices ternaires et  $\rho$  une racine cubique primitive de l'unité [...], en prenant pour les nonions fondamentaux  $u$  (l'unité ternaire)



$$\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

et les huit matrices  $m, m^2; n, n^2; m^2n; mn; m^2n^2$  construites avec les valeurs les plus simples de  $m, n$  qui satisfont aux équations

$$nm = pmn, m^3 = u, n^3 = u.$$

Les valeurs

$$m = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho^2 \end{vmatrix} \text{ et } n = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & \rho \\ \rho^2 & 0 & 0 \end{vmatrix}$$

[Sylvester, 1884c, 1337].

La construction des nonions nécessite la résolution d'"équations en matrices" du type  $nm = pmn$ . Ces équations amènent des problématiques nouvelles. Sylvester estime initialement que de telles équations impliquent une relation fonctionnelle entre les matrices  $m$  et  $n$ . Cette condition n'est cependant pas nécessaire en cas d'occurrence de racines latentes multiples :

Nommons, pour le moment,  $mn = u, nm = v$  ; on aura, comme auparavant  $uv = vu$  [...], on trouvera le moyen d'établir qu'en général cette équation amène à la conclusion que ou  $u$  doit être un *scalar*, c'est-à-dire de la forme

$$\begin{matrix} C & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & C \end{matrix},$$

ou bien  $v$  un *scalar*, ou sinon que  $nm, mn$  doivent être fonctions l'un de l'autre : mais on remarquera (ce qui m'avait échappé) que, si  $Fu = 0$  est l'équation identique en  $u$  et que la dérivée fonctionnelle  $F'u$  est une matrice vide (*vacuous*), c'est-à-dire dont le déterminant est zéro, le raisonnement est en défaut, cette vacuité a lieu dans le cas, et seulement dans le cas où deux des racines latentes (*lambdaïques*) de  $m$  sont égales. [...] Par exemple, si l'on fait

$$u = \begin{vmatrix} 0 & \rho & \rho^2 \\ 1 & 0 & 1 \\ \rho^2 & \rho & 0 \end{vmatrix}, v = \begin{vmatrix} 0 & 1 & 1 \\ \rho & 0 & \rho^2 \\ \rho & \rho^2 & 0 \end{vmatrix},$$

on trouvera

$$uv = \begin{vmatrix} -\rho & \rho & 1 \\ \rho & -\rho & 1 \\ \rho^2 & \rho^2 & -\rho \end{vmatrix} = vu$$

Mais on peut démontrer sans difficulté que  $v$  ne peut pas s'exprimer comme somme de puissances de  $u$ , ni *vice versa*  $v$  comme somme de puissances de  $u$ . [1884e, 274].

## • Les matrices dérogoires.

Sylvester interprète d'abord le cas où les solutions  $y$  et  $x$  de l'équation  $xy = yx$  ne seraient pas liées par une fonction rationnelle comme revenant à l'occurrence de racines latentes multiples. Dans un second temps, Sylvester déplace la question à l'étude de l'équation algébrique dont les racines latentes sont issues :

Avant de considérer l'équation  $xy = yx$ , il importe d'avoir une idée nette d'une certaine classe de matrices que je nomme *privilegiées* ou *dérogatoires*, en tant qu'elles dérogent à la loi générale que toute matrice est assujettie à satisfaire à une équation identique dont le degré ne peut pas être moindre que l'ordre de la matrice. Les matrices dérogoires sont justement celles qui satisfont à une équation d'un ordre inférieur à leur ordre propre; on peut les nommer *simplement, doublement, triplement, ... dérogoires*, selon que le degré de l'équation identique à laquelle elles satisfont diffère par une, deux, trois, ... unités du degré minimum ordinaire. [1884f, 471].



Le problème est ainsi déplacé de l'occurrence de racines multiples à l'examen des matrices dérogeant à la règle de satisfaire une équation de degré identique à leur ordre <sup>(28)</sup>. Si les matrices  $x$  et  $y$  sont non dérogoires, la relation  $xy = yx$  impliquera l'existence d'une relation fonctionnelle entre  $x$  et  $y$  même en cas d'occurrence de racines multiples:

En réservant les détails du calcul, voici le résultat général que j'ai démontré rigoureusement (en m'aidant de la notation des nonions) pour les matrices du troisième degré qui satisfont à l'équation  $xy = yx$ .

A moins que  $x$  ne soit une matrice privilégiée ou dérogoire,  $y$  sera toujours une fonction rationnelle et entière quadratique de  $x$ , et de même, à moins que  $y$  ne soit privilégiée,  $x$  sera une fonction pareille de  $y$ . [...] Il est bon de remarquer que nulle matrice ne peut être dérogoire, sauf pour le cas où il existe des égalités entre ses racines latentes ; mais ces égalités peuvent parfaitement subsister sans que la matrice à laquelle elles appartiennent soit dérogoire. [...] Il est à peine nécessaire d'ajouter que rien n'empêche, dans le cas où l'un ou l'autre de  $x$  et  $y$  où tous les deux sont dérogoires, qu'on puisse satisfaire à  $xy = yx$ , en supposant que  $x$  et  $y$  soient des fonctions explicites chacune l'une de l'autre : tout ce qu'on affirme, c'est que, dans le cas admis, cette supposition cesse d'être obligatoire ; c'est un cas très semblable à ce qui arrive dans le cas de défaut (failing case) du théorème de Maclaurin ; c'est celui où une variable est une fonction sans pouvoir être développée dans une série de puissance d'une autre variable. [1884f, 473].

Les conclusions de Sylvester manifestent une nouvelle lecture du mémoire de Cayley de 1858. C'est désormais le théorème de l'"équation identique" de Cayley qui fait l'intérêt principal de ce qui n'était jusqu'alors qu'une équation permettant de déterminer les racines latentes. A partir de 1884, une matrice devient, chez Sylvester, non seulement une *quantité*, mais une *quantité* qui vérifie une équation *algébrique*, "l'équation identique". La caractérisation des matrices dont l'équation identique est de degré moindre que l'équation caractéristique est l'objet principal des travaux de Sylvester à partir de 1884.

## 2. La formation des "espèces de matrices" d'Eduard Weyr.

Sur le continent, le premier mathématicien à employer la notion de matrice en référence à Cayley est un géomètre de Prague dénommé Eduard Weyr (quelques éléments biographiques sur Eduard Weyr sont donnés en [encart 8](#)). Les premiers travaux de Weyr sur les matrices sont inspirés des publications faites par Sylvester dans les *Compte Rendus* entre 1882 et 1885 et dont il a été question au paragraphe précédent. Dès 1884, Weyr s'inspire de la méthode élaborée par Sylvester pour exprimer des fonctions rationnelles de matrices afin d'étudier la fonction exponentielle « en supposant que les arguments soient des quaternions ». Mais c'est surtout la notion de matrice dérogoire de Sylvester qui va inspirer les travaux du géomètre de Prague par un rapprochement avec le problème de la caractérisation des *substitutions semblables*. Dans une note aux *Comptes Rendus* de 1885, Weyr emploie la notion de matrice dérogoire de Sylvester afin de "rectifier" la *caractérisation* des substitutions linéaires énoncée par Riemann en 1857 et basée sur la "décomposition" d'une substitution  $A$  en un produit de trois substitutions de la forme :

<sup>28</sup> En termes contemporains, il s'agit de déterminer le commutant d'un endomorphisme  $f$  d'un espace  $E$  de dimension finie : c'est-à-dire le sous espace  $\Gamma_f$  de  $E$  défini par  $\Gamma_f = \{g \text{ de } L(E) / f \circ g = g \circ f\}$ . Si  $f$  est diagonalisable,  $g$  appartient à  $\Gamma_f$  si et seulement si chaque sous espace propre  $E_\lambda$  de  $f$  est stable par  $g$ . Les matrices des  $g$  sont donc des matrices par blocs de tailles  $\dim E_\lambda$  et  $\dim \Gamma_f = \sum \dim E_\lambda^2$ . Si de plus les valeurs propres de  $f$  sont toutes distinctes, alors  $E_\lambda = \text{vect}(x_\lambda)$  est stable par  $g$  de  $\Gamma_f$  ; il existe  $\mu_\lambda$  tel que  $g(x_\lambda) = \mu_\lambda x_\lambda$ . La théorie des polynômes d'interpolations de Lagrange assure l'existence d'un polynôme  $P$  tel que  $P(\lambda_i) = \mu_i$  et  $\Gamma_f = \{P(f) / P \text{ de } K[X]\}$ . Le même polynôme peut être obtenu si le polynôme minimal  $\Pi_f$  est de degré  $n$ .

$$(A) = (\alpha) \left\{ \begin{array}{cccc} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{array} \right\} (\alpha)^{-1}$$

En cas de racines latentes multiples, la décomposition de Riemann est incorrecte : on ne peut pas, en général, caractériser la substitution  $A$  par une substitution semblable de la forme <sup>(29)</sup>:

$$\left\{ \begin{array}{cccc} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{array} \right\}$$

Nous avons vu que la notion de matrice avait permis à Sylvester de résoudre les difficultés posées par l'occurrence de *racines latentes multiples*, d'abord pour le problème des intersections de coniques en 1851 puis pour celui de l'expression des fonctions rationnelles d'une homographie en 1882. En 1885, Eduard Weyr investit les travaux de Sylvester et notamment les notions de « nullité » et de « matrice dérogoire » pour caractériser les substitutions linéaires en cas de racines multiples :

Si l'on a

$$\alpha = \alpha_1, \beta = \beta_1, \dots, \lambda = \lambda_1,$$

ce qui arrive, par exemple, quand les racines latentes sont toutes distinctes, on peut mettre  $M$  sous la forme

$$M = A^{-1} M_0 A,$$

$M_0$  étant une matrice dont la diagonale principale contient  $\alpha$  termes  $\mu_\alpha$ ,  $\beta$  termes  $\mu_\beta$ , ..., enfin  $\lambda$  termes  $\mu_\lambda$  et dont les autres termes sont nuls, et  $A$  désignant une matrice de nullité zéro; et ce n'est que dans le cas de

$$\alpha = \alpha_1, \beta = \beta_1, \dots, \lambda = \lambda_1,$$

qu'on peut mettre  $M$  sous cette forme. [Weyr, 1885a, 788].

Le problème de la caractérisation d'une substitution linéaire  $M$  de racines latente multiple • se ramène à l'étude de la matrice dérogoire  $M$  • dont Eduard Weyr donne une caractérisation basée sur l'expression de l'équation de degré minimal vérifiée par la matrice  $M$  • :

On sait que toute matrice de l'ordre  $n$  satisfait à une équation de degré  $n$  ; c'est l'équation fondamentale de M. Cayley. Il y a cependant des matrices qui satisfont à une équation de degré moindre que  $n$  : ce sont les matrices que M. Sylvester nomme dérogoires. Je suis parvenu à établir un théorème qui jette du jour sur ce sujet, et que je me permets de communiquer à l'Académie.

[...]

$M$  étant une matrice d'ordre  $n$  aux racines latentes

$$\mu_\alpha, \mu_\beta, \dots, \mu_\gamma$$

et

$$\alpha, \beta, \dots, \lambda$$

étant des degrés de multiplicité de ces racines, soient

$$\alpha_1, \beta_1, \dots, \lambda_1$$

les degrés de nullité des matrices

$$M_{-\mu_\alpha}, M_{-\mu_\beta}, \dots, M_{-\mu_\gamma};$$

alors  $M$  satisfait à l'équation

$$(M - \mu_\alpha)^{\alpha - \alpha_1 + 1} (M - \mu_\beta)^{\beta - \beta_1 + 1} \dots (M - \mu_\lambda)^{\lambda - \lambda_1 + 1} = 0$$

<sup>29</sup> En termes contemporains, il est nécessaire de recourir à la forme de Jordan ou à la forme canonique rationnelle dont il a été question dans la première partie.

Les nombres

$$\alpha_1, \beta_1, \dots, \lambda_1$$

dont chacun est au moins égal à 1, ne peuvent pas surpasser les nombres respectifs

$$\alpha, \beta, \dots, \lambda.$$

Dans le cas de

$$\alpha_1 = \beta_1 = \dots = \lambda_1 = 1,$$

on tombe sur l'équation de M. Cayley. Dans tout autre cas, la matrice M est dérogoire. [1885a, 788].

Ces considérations sont ensuite appliquées au problème de la caractérisation des substitutions linéaires, formulé comme relevant d'une "répartition des matrices en espèces", et résolu par l'introduction d'un système d'invariants, la « caractéristique » :

Je dis, de deux matrices d'ordre  $n$ , qu'elles sont de même espèce si elles possèdent les mêmes racines aux mêmes caractéristiques.

$M$  et  $N$  étant deux matrices de même espèce, on peut toujours assigner des matrices  $Q$ , de nullité zéro, telles qu'on ait

$$N = Q^{-1}MQ.$$

Et, réciproquement, deux matrices  $M$  et  $N$ , liées par une telle équation, sont de même espèce.

[...] Soient  $M$  une matrice quelconque d'ordre  $n$  et  $\mu_\alpha$  une racine  $\alpha^{\text{uple}}$  de  $M$ . En formant les puissances de  $M - \mu_\alpha$ , on tombe nécessairement sur une puissance  $(M - \mu_\alpha)^\rho$  qui est de nullité  $\alpha$ ; les puissances plus élevées sont de la même nullité. [...] Désignons par

$$\alpha_1, \alpha_1 + \alpha_2, \dots, \alpha_1 + \alpha_2 + \dots + \alpha_p = \alpha$$

les degrés de nullité des matrices

$$M - \mu_\alpha, (M - \mu_\alpha)^2, \dots, (M - \mu_\alpha)^\rho;$$

$$[\dots] \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_p.$$

Pour abrégé, je dis que la racine  $\mu_\alpha$  a pour *caractéristiques* les nombres

$$(\alpha, \alpha_1, \alpha_2, \dots, \alpha_p).$$

[Weyr, 1885b, 966].

Deux matrices de même espèce ont même caractéristique. Pour démontrer la réciproque de cette propriété, Weyr articule la *forme* matricielle à l'*invariant* qu'est la caractéristique :

Les entiers

$$\alpha, \alpha_1, \dots, \alpha_p; \beta, \beta_1, \dots, \beta_\sigma; \dots; \lambda, \lambda_1, \dots, \lambda_\tau$$

ayant été choisis de manière que chacun d'eux soit au moins égal à 1, et que les suites

$$(\alpha_1, \dots, \alpha_p), (\beta_1, \dots, \beta_\sigma), \dots, (\lambda_1, \dots, \lambda_\tau),$$

ne soient jamais croissantes, et que, de plus,

$$\alpha = \alpha_1 + \dots + \alpha_p, \beta = \beta_1 + \dots + \beta_\sigma, \dots, \lambda = \lambda_1 + \dots + \lambda_\tau,$$

$$n = \alpha + \beta + \dots + \lambda,$$

je dis qu'il existe toujours des matrices d'ordre  $n$ , ayant les racines  $\mu_\alpha, \mu_\beta, \dots, \mu_\lambda$  aux caractéristiques respectives

$$(\alpha, \alpha_1, \dots, \alpha_p), (\beta, \beta_1, \dots, \beta_\sigma), \dots, (\lambda, \lambda_1, \dots, \lambda_\tau),$$

les valeurs  $\mu_\alpha, \mu_\beta, \dots, \mu_\lambda$  étant arbitraires, mais distinctes entre elles. [Weyr, 1885, 966].

La démonstration est basée sur la construction par itération d'une matrice d'ordre  $\alpha$  ayant une unique racine  $\mu_\alpha$  de multiplicité  $\alpha$  et de caractéristique  $(\alpha, \alpha_1, \dots, \alpha_p)$ . En cas d'occurrence de racines multiples, le résultat n'est pas toujours une forme diagonale mais Weyr parvient à une décomposition de la matrice en articulant la décomposition de la *forme* polynomiale de l'équation minimale à une décomposition en "compartiments" de la *forme* de la matrice :

Pour cet effet, désignons par  $G_{p-1-\mu_\alpha}$  la matrice zéro, et d'ordre  $\alpha_p$ , et posons successivement

$$G_{\rho-2-\mu\alpha} = \left\{ \begin{array}{c|c} \frac{G_{\rho-1-\mu\alpha}}{A_{\rho-1}} & \overbrace{\begin{pmatrix} 0 & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & 0 \end{pmatrix}}^{(\alpha_{\rho-1})} \end{array} \right\}, G_{\rho-3-\mu\alpha} = \left\{ \begin{array}{c|c} \frac{G_{\rho-2-\mu\alpha}}{A_{\rho-2}} & \overbrace{\begin{pmatrix} 0 & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & 0 \end{pmatrix}}^{(\alpha_{\rho-2})} \end{array} \right\},$$

$$G_{1-\mu\alpha} = \left\{ \begin{array}{c|c} \frac{G_{2-\mu\alpha}}{A_2} & \overbrace{\begin{pmatrix} 0 & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & 0 \end{pmatrix}}^{(\alpha_2)} \end{array} \right\}, H_{-\mu\alpha} = \left\{ \begin{array}{c|c} \frac{G_{1-\mu\alpha}}{A_1} & \overbrace{\begin{pmatrix} 0 & \dots & 0 \\ \cdot & \dots & \cdot \\ 0 & \dots & 0 \end{pmatrix}}^{(\alpha_1)} \end{array} \right\}.$$

[...] Les compartiments  $A_{\rho-1}, A_{\rho-2}, \dots, A_1$  sont formés de la manière suivante :

$$A_{\rho-1} = \left\{ \begin{array}{c|c} \overbrace{\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix}}^{(\alpha_\rho)} & (\alpha_{\rho-1}) \end{array} \right\}, A_{\rho-2} = \left\{ \begin{array}{c|c} \overbrace{\begin{pmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \cdot & \dots & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix}}^{(\alpha_\rho)} & (\alpha_{\rho-1}) \end{array} \right\},$$

Dans le cas de  $\alpha_\rho = \alpha_{\rho-1}$ , le compartiment  $A_{\rho-1}$  aura la forme d'un carré et ne contiendra pas les lignes remplies entièrement de zéros; les mêmes lignes manqueront dans  $A_{\rho-2}$ , si  $\alpha_{\rho-1} = \alpha_{\rho-2}$ , et ainsi de suite. A l'aide de  $H$  on peut former une matrice  $K$  d'ordre  $\alpha + \beta$  ayant les racines  $\mu_\alpha$  et  $\mu_\beta$  aux caractéristiques  $(\alpha, \alpha_1, \dots, \alpha_\rho), (\beta, \beta_1, \dots, \beta_\sigma)$ . [Weyr, 1885b, 967].

Dans la méthode de "décomposition" des matrices de Weyr, le *calcul symbolique* et l'*extraction de mineurs* de la représentation matricielle sont indissociables. Un calcul de *puissances* de matrices donne une *décomposition* en *compartiments* qui reflète la *décomposition polynomiale* de l'invariant qu'est le polynôme minimal. Associant des pratiques issues de réseaux différents dont certaines remontent aux années 1850 comme le calcul des matrices de Cayley, Weyr développe une *combinatoire* sur la représentation matricielle par un processus d'*itération* du calcul des *puissances successives* de "compartiments" d'une matrice. Cette approche deviendra, dans les années 1920-1930, la méthode la plus répandue de démonstration du théorème de Jordan de la décomposition matricielle ([encart 9](#)) <sup>(30)</sup>.

### 3. La rencontre de la théorie des formes bilinéaires et de la théorie des matrices.

<sup>30</sup> Les recherches de Weyr sont indépendantes de la forme canonique énoncée par Jordan en 1870, sur le rôle de Jordan dans l'histoire du théorème de Jordan de la décomposition matricielle, voir [Brechenmacher, 2006].

En termes contemporains, Weyr introduit la suite des dimensions des sous espaces caractéristiques d'un endomorphisme sur un espace vectoriel de dimension finie. Soit  $f \in L(E)$  et  $\lambda$  une valeur propre de  $f$ , il existe un unique entier naturel  $r$  tel que :

$$\{0\} = \text{Ker}(f-\lambda I)^0 \subsetneq \text{Ker}(f-\lambda I) \subsetneq \dots \subsetneq \text{Ker}(f-\lambda I)^r = \text{Ker}(f-\lambda I)^{r+1}$$

Entre 1885 et 1890, Eduard Weyr est le premier et le seul mathématicien du continent à employer la représentation matricielle. En 1890, il élabore une *synthèse théorique* qui mêle les *notions auparavant distinctes de formes bilinéaires et de matrices*. Le mémoire intitulé "Sur les formes bilinéaires", publié par Weyr dans le premier numéro des *Monatsberichte für Mathematik und Physik*, a pour objet de réorganiser la théorie des formes bilinéaires par la notion « plus abstraite » de matrice :

L'objectif des considérations qui suivent [...] est d'introduire un nouveau moyen d'action dans la théorie des formes bilinéaires par la considération de systèmes de valeurs associés à une matrice, cette méthode permet la solution de plusieurs problèmes et, en particulier, du problème de la transformation simultanée de deux formes bilinéaires, résolu par Weierstrass.

[...] Une forme bilinéaire

$$\sum_{(h,k)} a_{hk} x_h y_k, (h, k = 1, 2, \dots, n)$$

est parfaitement déterminée par une matrice du  $n^e$  ordre  $\{a_{hk}\}$ , c'est à dire par un système de  $nn$  coefficients ordonnés en  $n$  lignes et  $n$  colonnes ;[...] Dans la suite nous considérerons le concept de forme bilinéaire avec celui, plus abstrait, de matrice. [Weyr, 1890, 163, traduction, F.B.].

A la fin du XIX<sup>e</sup> siècle, la théorie des formes bilinéaires jouait un rôle essentiel dans de nombreuses branches des mathématiques allant de l'arithmétique aux systèmes différentiels en passant par la théorie des groupes. Le rôle joué par la notion de forme bilinéaire vers 1890 est, pour simplifier, analogue à celui que jouera la notion de matrice après les années 1930. Par exemple, le problème de la résolution des systèmes d'équations différentielles linéaires était considéré comme appartenant à la théorie des formes bilinéaires, un système du second ordre se représentant sous la forme :

$$(I) \quad \frac{d^2}{dt^2} \frac{dA}{dx_1} = \frac{dB}{dx_1}, \dots, \frac{d^2}{dt^2} \frac{dB}{dx_n} = \frac{dB}{dx_n}$$

où  $A = \sum a_{hk} x_h y_k$ ,  $B = \sum b_{hk} x_h y_k$

Les manipulations des variables par lesquelles on transforme le système différentiel reviennent à *transformer* le couple de formes  $(A, B)$ . Le problème central de la théorie des formes bilinéaires est de déterminer, deux *couples* de formes  $(A, B)$  et  $(A', B')$  étant donnés, si l'on peut transformer l'un en l'autre. Il s'agit donc de caractériser des *classes d'équivalences des formes ou des couples de formes*. Dans le cadre de la théorie établie par Frobenius en 1878-1879, la transformation linéaire des formes est traitée comme une opération de multiplication portant sur des "systèmes de  $n^2$  valeurs".

Si l'on applique une substitution sur une seule suite de variables d'une forme bilinéaire, on obtient de nouveaux coefficients qui définissent une forme transformée comme si la substitution était une opération entreprise sur la forme elle-même. Il apparaît donc que la distinction entre *operandus* et *opérateur* s'efface dans le résultat de la même manière que, dans le cas de la multiplication, *multiplicandus* et *multiplicateurs* sont confondus ou encore de la même façon que, dans le calcul des quaternions, les systèmes de coordonnées de l'espace se confondent avec les opérations sur ces systèmes. Ces considérations m'ont conduit traiter la transformation des formes bilinéaires comme une composition de substitutions linéaires. [...] Je vais ici multiplier une forme par une constante, additionner deux formes et différentier une forme dont les coefficients dépendent d'un paramètre. Je ne vais cependant pas multiplier deux formes l'une par l'autre. Il ne doit donc pas y avoir d'ambiguïté, lorsque je désigne la forme  $P$  qui met ensemble  $A$  et  $B$  comme le produit des formes  $A$  et  $B$ , qui seront appelées des facteurs de  $P$ . [...] On dit qu'une forme  $B$  est équivalente à la forme  $A$ , quand il existe deux formes  $P$  et  $Q$  de déterminants non nuls vérifiant l'équation  $PAQ = B$  [...].  $P$  et  $Q$  se nomment les substitutions par lesquelles  $A$  se transforme en  $B$ . Toutes les formes équivalentes forment une *classe* de forme. [Frobenius, 1878, 343 ; traduction, F.B.].

Le "produit" de deux formes bilinéaires  $A$  et  $B$ ,  $P = \sum \frac{dA}{dy_\gamma} \cdot \frac{dB}{dx_\gamma}$ , est une forme dont les coefficients "mêlent ensemble" ceux de  $A$  et  $B$ . Le calcul symbolique permet de représenter les relations arithmétiques d'équivalence des formes, la transformation de la forme  $A$  par les substitutions linéaires  $x_\alpha = \sum p_{\alpha\beta} X_\beta$  et  $y_\alpha = \sum q_{\alpha\beta} Y_\beta$  s'exprimant comme un produit symbolique de trois formes  $P'AQ$  où  $P = \sum p_{\alpha\beta} x_\alpha y_\beta$  et  $Q = \sum q_{\alpha\beta} x_\alpha y_\beta$  (<sup>31</sup>). Exprimés de manière symbolique, le problème de l'équivalence des couples de formes  $P(rE-A)Q = rE-B$ , d'une part, et celui de la similitude des substitutions  $P^{-1}AP=B$ , d'autre part, sont susceptibles d'une même approche et résolus par la définition d'un système complet d'invariants polynomiaux, les diviseurs élémentaires et facteurs invariants inspirés d'un théorème énoncé par Weierstrass en 1868 ([encart 10](#)).

Dans le cadre de travaux menés dans les années 1970-1980 sur l'histoire de la théorie des systèmes hypercomplexes et des algèbres associatives, Karen Parshall et Thomas Hawkins identifiaient la théorie des formes bilinéaires de Frobenius à la théorie des matrices de Cayley. Aux yeux des deux historiens l'utilisation, dans les deux cas, d'un calcul symbolique sur des systèmes de nombres et l'énoncé de deux théorèmes considérés comme identiques (le théorème de Cayley Hamilton) justifiait de considérer le lien entre les notions de matrices et de formes bilinéaires comme non problématique et même préexistant à son énoncé par les mathématiciens. De manière similaire, Jean Dieudonné écrivait en 1977 :

L'étude plus approfondie des matrices, et notamment de leur classification sous divers point de vue, allait se faire par le truchement d'une théorie équivalente, celle des formes bilinéaires. [...] Le rôle de législateur qu'avait joué Cauchy pour la théorie des déterminants est tenu par Frobenius pour la théorie des matrices. Dans plusieurs mémoires publiés entre 1877 et 1880, il reprend la plupart des résultats précédents en les développant d'une manière systématique et y ajoutant de nombreux compléments. [Dieudonné, 1977, 97-99].

Or Frobenius, contrairement à ce qui est souvent affirmé, connaissait le mémoire de Cayley en 1878 et c'est sciemment qu'il n'adoptait pas la notion de matrice dans sa théorie des formes bilinéaires (<sup>32</sup>). D'une part, avant le mémoire publié par Weyr en 1890, aucune relation n'avait été faite entre les théories de Cayley et de Frobenius, d'autre part, dans les années qui suivront la publication de Weyr, les matrices envahiront les textes mathématiques publiés en Allemagne. Comprendre la nouvelle popularité des matrices sur le continent après 1890 nécessite de suivre une approche complémentaire aux travaux de Thomas Hawkins et Karen Parshall, en s'attachant à étudier l'évolution qu'implique, pour la notion de matrice, sa rencontre avec la théorie des formes bilinéaires. .

Le mémoire publié par Weyr en 1890 est conçu en deux parties, la première développe une théorie du "concept abstrait" de matrice que la seconde "applique" aux problèmes classiques de la théorie des formes bilinéaires de Frobenius comme la caractérisation des formes semblables et l'équivalence des couples de formes. Par sa rencontre avec la théorie des formes bilinéaires par laquelle, pour reprendre une expression de Sylvester, "une matrice se fait dérober ses dimensions spatiales et représentée comme une somme linéaire", la notion de matrice évolue. Dans la réorganisation de la théorie des formes bilinéaires proposée par Weyr, en 1890, une matrice n'a plus d'existence autonome, elle est associée à un système de valeurs et à la notion essentielle d'indépendance linéaire des systèmes. Le nombre maximum de systèmes linéairement indépendants que l'on peut extraire d'un système de  $n$  valeurs données

<sup>31</sup> La notation  $P'$  désigne la transposée  $'P$ . Sur la pratique algébrique de Frobenius voir [Brechenmacher, 2006, 247-271].

<sup>32</sup> Même en Angleterre, la notion de matrice est peu utilisée et son emploi par quelques mathématiciens comme Smith en 1861 est sans commune mesure avec la popularité des matrices dans les années 1890.



est un invariant, dénommé *rang*  $r$  du système en référence aux travaux menés par Kronecker dans les années 1880 sur la théorie arithmétique des grandeurs algébriques<sup>(33)</sup>. Le rang est un invariant complémentaire de la nullité de Sylvester, il correspond à l'ordre du plus grand mineur non nul du déterminant  $|A|$  et si  $v$  est la nullité de  $A$  alors  $r+v=n$ . La nullité d'une matrice  $M$  est  $v$  si et seulement si il existe  $v$  systèmes (et pas plus) indépendants  $(x_i)$  tels que  $A(x_i)=0$ <sup>(34)</sup>. Cette association de la nullité des matrices et de la notion de rang caractérisant les systèmes linéairement indépendants sous tend une *pratique algébrique* dont Weyr fait un usage constant dans son mémoire de 1890. Elle permet notamment de démontrer que si  $A$  est de nullité  $\alpha$ ,  $M$  de nullité  $\mu$ , alors le produit  $P=AM$  est de nullité  $\omega$  telle que  $\omega \geq \alpha$ ,  $\omega \geq \mu$  et  $\mu + \alpha \geq \omega$ . Cette propriété permet d'encadrer la nullité d'un polynôme de matrices décomposé en facteurs linéaires<sup>(35)</sup>, les nullités des puissances d'une matrice forment alors une suite croissante :

Soit  $M$  une matrice d'ordre  $n$ , dont la racine 0 à pour multiplicité  $\alpha$  et pour nullité  $\alpha_1$  alors,  
 1. soit  $\alpha_1=\alpha$  et toutes les puissances positives  $M^k$  ont pour nullité  $\alpha_1$  ;  
 2. soit  $\alpha_1 < \alpha$  et  $M^2$  a pour nullité  $\alpha_1 + \alpha_2$ , où  $\alpha_1 \geq \alpha_2 > 0$  et  $\alpha_2 \geq \alpha - \alpha_1$  où  $\alpha_2 = \alpha - \alpha_1$ . Si  $M^2$  a pour nullité  $\alpha$ , alors toutes les puissances successives  $M^k$  ( $k > 2$ ) ont pour nullité  $\alpha$ .  
 [Weyr, 1890, 186, Traduction, F.B.].

L'examen de la suite des puissances des matrices  $M-\mu$ , apporte un éclairage sur les relations entre la multiplicité d'une racine caractéristique  $\mu$ , le nombre de solutions de l'équation  $M(x)=\mu x$  (égal au rang de la matrice  $M-\mu_a$ ), le polynôme minimal de  $M$  et son polynôme caractéristique:

Il s'en suit que les expressions des puissances

$$M, M^2, M^3, \dots$$

s'achèvent à une puissance donnée  $M^p$  de nullité  $\alpha$  et que toutes les puissances supérieures ont pour nullité  $\alpha$ . Désignons par les nombres

$$\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3, \dots, \alpha_1 + \alpha_2 + \dots + \alpha_p = \alpha$$

les nullités des matrices  $M, M^2, M^3, \dots, M^p$ , on a alors

$$\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \dots \geq \alpha_p > 0$$

Si la racine  $\mu_a$  de la matrice  $M$  a pour multiplicité  $\alpha$ , alors, d'après l'Art. 15, la matrice  $M-\mu_a$  a pour racine 0 à la multiplicité  $\alpha$ . Soit alors

$$\alpha_1, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3, \dots, \alpha_1 + \alpha_2 + \dots + \alpha_p = \alpha$$

les nullités des matrices  $M-\mu_a, (M-\mu_a)^2, (M-\mu_a)^3, \dots, (M-\mu_a)^p$ , on appelle nombres caractéristiques de la racine  $\mu_a$  les nombres  $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_p$ . [...]

Soit  $\alpha_1, \alpha_2, \dots, \alpha_p$  les nombres caractéristiques de la racine  $\mu_a$ , et de même

$\beta_1, \beta_2, \dots, \beta_\sigma$  ceux de la racine  $\mu_\beta, \dots, \lambda_1, \lambda_2, \dots, \lambda_\tau$  ceux de  $\mu_\lambda$ . Alors  $(M-\mu_a)^\rho$  a pour nullité  $\alpha$ ,  $(M-\mu_\beta)^\sigma$  a pour nullité  $\beta$ , ..., et enfin  $(M-\mu_\lambda)^\tau$  a pour nullité  $\alpha + \beta + \dots + \lambda$ . On a donc

$$(M-\mu_a)^\rho (M-\mu_\beta)^\sigma \dots (M-\mu_\lambda)^\tau = 0.$$

et cette équation est l'équation de plus bas degré à coefficients scalaires vérifiée par la matrice  $M$  ; elle sera dénommée équation fondamentale de  $M$ . [...]

$$f(M)=0$$

Il s'agit là de l'équation énoncée par Cayley et que chaque matrice d'ordre  $n$  vérifie ; si  $M$  a pour ensemble de racines  $\mu_a, \mu_\beta, \dots, \mu_\lambda$ , de multiplicités  $\alpha, \beta, \dots, \lambda$ , alors le polynôme

$$f(M) = (-1)^n (M-\mu_a)^\alpha (M-\mu_\beta)^\beta \dots (M-\mu_\lambda)^\lambda$$

a pour diviseur  $\varphi(M)$ , et il s'ensuit de (6) que l'équation de Cayley est identique à l'équation fondamentale d'une matrice si et seulement si

$$\rho = \alpha, \sigma = \beta, \dots, \tau = \lambda.$$

[Weyr, 1890, 186-188, traduction, F.B.].

<sup>33</sup> Au sujet des travaux arithmétiques de Kronecker, voir [Brechenmacher, 2006].

<sup>34</sup> Ce qu'on appellerait la dimension du noyau d'un opérateur.

<sup>35</sup> Weyr définit d'abord la suite croissante des nullités des puissances des matrices, puis d'un produit de deux polynômes premiers entre eux de matrices (Weyr démontre que la nullité d'un produit de deux polynômes premiers entre eux est le produit des nullités de chaque terme du produit).



Si la matrice  $M$  a pour nullité  $\alpha_l$ , il existe  $\alpha_l$  systèmes indépendants  $(x^{(\alpha_l+1)}), \dots, (x^{(n)})$  vérifiant l'équation  $M(x)=0$ . Ces systèmes peuvent être complétés en  $n$  systèmes indépendants formant une nouvelle matrice  $\{(x'), (x''), \dots, (x^{(n)})\}$ . Le produit  $M\{(x'), (x''), \dots, (x^{(n)})\}$  est une matrice dont les  $\alpha_l$  premières colonnes sont nulles et la matrice

$$M_l = \{(x'), (x''), \dots, (x^{(n)})\}^{-1} M \{(x'), (x''), \dots, (x^{(n)})\}$$

a la forme :

$$M_l = \left\{ \begin{array}{ccc|c} 0 & \dots & 0 & S \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & G \end{array} \right\}$$

Weyr considère alors successivement les matrices :

$$M_2^1 = \left\{ \begin{array}{ccc|c} 0 & \dots & 0 & S_1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & G^2 \end{array} \right\} \text{ de nullité } \alpha_l + \text{null. } G$$

$$M_1^3 = \left\{ \begin{array}{ccc|c} 0 & \dots & 0 & S_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & G^3 \end{array} \right\}$$

La suite des nullités des puissances de  $M$  est donc une suite de nombres entiers croissante et bornée par la multiplicité de la racine caractéristique 0. La méthode d'itération de Weyr permet d'obtenir une décomposition matricielle quelle que soit la multiplicité des racines caractéristiques de la matrice initiale. En s'appuyant sur ses travaux de 1885 sur la « répartition des matrices en espèces », Eduard Weyr va, en 1890, proposer un véritable retournement théorique dans la théorie des formes bilinéaires. La théorie qu'il élabore n'est plus centrée sur des énoncés d'invariants polynomiaux mais met en avant *le caractère opératoire de la représentation matricielle afin de la décomposer à des formes canoniques*.

## Conclusion.

En parcourant, dans les parties II et III de cet article, des textes publiés entre 1850 et 1890 nous avons suivi des *évolutions* de la notion de matrice dans le temps. Entre sa première définition comme *mère des mineurs* d'un déterminant par Sylvester en 1851 et la décomposition matricielle de Weyr, *les matrices supportent des significations et des pratiques multiples*. Dans les années 1850 les matrices *mère des mineurs* de Sylvester et les *matrices* de Cayley étaient des notions distinctes et supportaient les pratiques différentes d'*extractions de mineurs* d'un déterminant d'une part, de *calcul symbolique* d'autre part. Elles renvoyaient également à des contextes différents, les matrices de Sylvester étaient attachées au problème de la *multiplicité* des racines d'une équation obtenue par un calcul déterminant, celles de Cayley visaient l'*expression polynomiale* de *fonctions rationnelles* d'homographies dans la tradition des travaux de l'*école algébrique anglaise*. En suivant le détail des travaux de Sylvester entre 1882 et 1885 nous avons vu la notion de matrice *évoluer* chez un même auteur. La synthèse théorique de Weyr en 1890, basée sur la rencontre de pratiques matricielles et de la théorie des formes bilinéaires nous a donné un exemple de la manière dont des *pratiques*, élaborées sur une longue période et dans des contextes divers, se *rencontrent* en provoquant un enrichissement du champ des significations. Nous pouvons à présent revenir sur la question de l'*identité* de la notion de matrice posée en introduction. *Y a-t-il une ou plusieurs notions de matrice dans l'histoire ?* Les matrices de Weyr (1890) et de Cayley (1890) sont-elles les *mêmes* ? Ce sont, d'une part, les mêmes selon Weyr lui-même qui, dès le premier paragraphe de son mémoire, attribue la notion de matrice à Cayley.

Les règles du calcul avec les *matrices* données par *Cayley* dans son exposé "A memoir on the Theory of Matrices", Philos. Transactions of the R. Society, London 1859, vol 148, s'appliquent directement pour la composition des *formes bilinéaires* (*Frobenius*, Ueber lineare Substitutionen und bilineare Formen, Journal für Mathematik. Bd. LXXXIV, §1 sqq) ; [...]. [Weyr, 1890, 163, traduction, F.B.].

Ce sont des notions différentes d'autre part, puisque la référence à Cayley se mêle à la citation d'autres travaux comme ceux de Frobenius (1878), de Kronecker (1884) ou de Sylvester (1884). Le mémoire de Weyr construit une *nouvelle identité* entre matrices et formes bilinéaires par laquelle la notion de matrice de Cayley, évolue, s'enrichit, change de signification.

Si l'identité de la notion de matrice évolue dans le temps, il faut poser la question des permanences et des héritages. Comment Weyr, par exemple, lit-il en 1890 le mémoire de Cayley de 1858 ? La comparaison du mémoire de Weyr et de celui de Cayley réalisée en [encart 12](#) met en évidence une *organisation parallèle des deux textes*. La structure des deux mémoires vise la démonstration de ce que Weyr désigne comme le "théorème fondamental de M. Cayley" selon lequel "une matrice satisfait une équation algébrique de son propre ordre". A cette similitude des deux textes, il faut opposer une différence essentielle. Tandis que Cayley définit les matrices comme des "quantités arrangées en forme de carré", la définition de Weyr est basée sur l'écriture d'une "équation symbolique" à la manière de Frobenius. *La caractéristique la plus reconnaissable de la matrice de Cayley, sa forme, sa représentation en "carré", ne joue donc aucun rôle dans la présentation du "calcul des matrices" par Weyr.* Quelle est alors l'héritage de Cayley ? En quoi la notion de matrice enrichit-elle la théorie des formes bilinéaires en 1890 ?

*Il est tout à fait significatif que, dans le mémoire de Weyr, la première représentation d'une matrice par un tableau n'intervienne qu'après une vingtaine de pages comme méthode de résolution d'un problème précis, problème que nous avons abordé dans la troisième partie de cet article et consistant à démontrer que l'invariant qu'est la "caractéristique" détermine une*

classe de matrices semblables <sup>(36)</sup>. Ce qui est propre à la théorie des matrices de Cayley est aussi l'aspect qui peut paraître le plus étrange au lecteur contemporain, *la notion de "single quantity"* ("Scalarmatrizen" chez Weyr) <sup>(37)</sup>. C'est d'ailleurs avec la définition de la notion de "matrice scalaire" que les analogies géométriques ("diagonales", "triangles", "carrés"), étrangères au calcul symbolique de Frobenius, apparaissent dans le texte d'Eduard Weyr :

La matrice nulle et la matrice unité sont manifestement des cas particuliers de ces matrices, dont les éléments de la diagonale principale possèdent la même valeur  $a$  : de telles matrices seront appelées matrices scalaires et désignées par le symbole  $a$ . Par conséquent les matrices  $I$  et  $0$  seront désignées par  $1$  et  $0$  respectivement. [Weyr, 1890, 165, traduction, F.B.].

Une *matrice scalaire* est à la fois un *nombre*  $a$  (un scalaire), et une matrice réduite à une *diagonale* composée de la succession d'une même valeur scalaire. Il s'agit donc d'une quantité à la fois simple et multiple. De cette dualité, Cayley, et Weyr à sa suite, élaborent une *méthode généralisant l'écriture polynomiale aux polynômes de matrices* du type  $a_0A^v + a_1A^{v-1} + \dots + a_v$ . Contrairement aux fonctions entières de formes employées par Frobenius dans la théorie des formes bilinéaires, les  $a_i$  ne sont pas des nombres mais des matrices scalaires qui permettent de factoriser les polynômes de matrices en facteurs linéaires sur le modèle des polynômes de nombres <sup>(38)</sup>:

Soit à présent  $A = \{a_{hk}\}$  une matrice quelconque [...]. La matrice

$$a_0A^v + a_1A^{v-1} + \dots + a_v,$$

où  $a_0, a_1, \dots, a_v$  désignent des matrices scalaires et  $v$  un nombre entier positif, sera dénommée fonction entière de la matrice  $A$ .

Si  $a_0, a_1, \dots, a_v$  et  $x$  sont des grandeurs complexes, on a pour tout  $x$  l'identité

$$a_0x^v + a_1x^{v-1} + \dots + a_v = a_0(x-\rho_1)\dots(x-\rho_v),$$

qui donne, par la considération de la règle de calcul précédente, également la relation

$$a_0A^v + a_1A^{v-1} + \dots + a_v = a_0(A-\rho_1)\dots(A-\rho_v),$$

où  $A$  désigne une matrice quelconque.

[Weyr, 1890, 165, traduction, F.B.].

La factorisation des équations de polynômes de matrices en "facteurs latents"  $M-\mu$ , que nous avons pu voir à l'œuvre dans la troisième partie chez Sylvester et chez Weyr, étend les méthodes polynomiales aux matrices. En mêlant des pratiques rattachées aux travaux de Sylvester et à la notion de matrice *mère des mineurs*, les pratiques polynomiales attachées à la « single quantity » de Cayley ainsi que des pratiques développés dans d'autres réseaux par des mathématiciens comme Kronecker et Frobenius, Weyr élabore des *procédés opératoires portant sur la représentation matricielle* qui permettent une *décomposition à des formes canoniques*.

Il faut donc reconnaître, parallèlement aux évolutions d'identité de la notion de matrice, un *héritage* dans les pratiques élaborées par les auteurs du réseau que nous avons considéré. Cet héritage renvoie à une *spécificité* mathématique des *pratiques matricielles* que nous allons mettre en évidence en comparant les méthodes de Weyr à celles d'un auteur de la même époque mais dont les recherches s'inscrivent dans un réseau différent : Jean Louis Sauvage. En 1890, Sauvage publie, sous le nom de "théorie des diviseurs élémentaires", le premier exposé en français de la théorie des formes bilinéaires de Frobenius. Il explique notamment comment obtenir, à partir de la composition polynomiale  $[P, Q] = \Pi_i [P_i, Q_i]$  des diviseurs

<sup>36</sup> Comme nous l'avons vu, chez Weyr, une matrice applique un "système de valeurs" sur un autre et, si Weyr utilise une notation symbolique  $(x)$  pour désigner les matrices, c'est, au-delà de l'influence de Frobenius déjà mentionnée, que la "représentation en carré" de Cayley est relative à la donnée d'un système de valeur et pose donc une difficulté de représentation.

<sup>37</sup> Cette notion est qualifiée d'"étrange" en référence à un article de Tony Crilly [1978] qui souligne toute l'ambiguïté pour les mathématiques contemporaines de l'emploi qu'en fait Cayley.

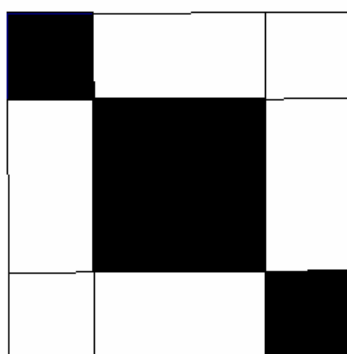
<sup>38</sup> Lorsque Frobenius recourt à une factorisation, par exemple pour exprimer les diviseurs élémentaires de Weierstrass, cette factorisation concerne le déterminant du polynôme de formes et non le polynôme de formes lui-même.

élémentaires du déterminant  $[P, Q]$  un couple canonique de formes  $(P, Q)$  par une *composition* des mineurs  $[P_i, Q_i]$ . Comme nous le voyons ci-dessous, Sauvage décrit la manière dont les mineurs  $[P_i, Q_i]$  se "composent" en un même déterminant  $[P, Q]$  en *articulant la composition polynomiale des diviseurs élémentaires à une composition géométrique en "carrés" et "rectangles" de la représentation par tableau du déterminant et de ses mineurs* :

40. Cherchons les déterminants d'ordre  $\hat{Q}$  de  $[P, Q]$ . Chacun d'eux correspond à une succession de  $\hat{Q}$  lignes et de  $\hat{Q}$  colonnes dans le déterminant principal. Celui-ci peut être représenté schématiquement par la fig. 1 composée :

- 1° de carrés noirs ayant une diagonale commune avec le carré principal ;
- 2° de carrés et de rectangles blancs (fig. 1).

Fig.1



Les parties blanches correspondent à des éléments tous nuls du déterminant  $[P, Q]$  et les carrés noirs correspondent aux déterminants  $[P_i, Q_i]$ . Cela posé, si l'on supprime  $\hat{Q}$  lignes et  $\hat{Q}$  colonnes de  $[P, Q]$ , on aura un mineur que l'on pourra représenter schématiquement par la fig. 2, analogue à la précédente.

Fig. 2



Mais, dans cette nouvelle figure, il pourra se présenter des *rectangles noirs*. Supposons que cette circonstance se produise, et, pour fixer les idées, imaginons que la première partie noire soit un rectangle renfermant  $k$  lignes et  $k'$  colonnes, et allongé dans le sens horizontal, c'est à dire que l'on a  $k' > k$ .

Dans un élément quelconque du déterminant représenté par la fig. 2 entreront forcément un élément non nul de la première ligne, un élément non nul de la seconde ligne...un élément non nul de la  $k^{\text{ème}}$  ligne, et à cause de la forme de la fig. 2, ces éléments devront appartenir à  $k$  des  $k'$  premières colonnes. Mais il restera encore  $k' - k$  des premières colonnes dans chacune desquelles il faudra prendre un élément, et cet élément, devant être pris en dehors des  $k$  premières lignes, appartiendra à une partie blanche de la fig. 2 et ne pourra être qu'un zéro. Nous en concluons que, dans le développement du mineur considéré, tous les éléments sont nuls et, par suite, que ce mineur lui-même est identiquement nul.

41. Nous sommes donc autorisés à ne considérer que les mineurs de  $[P, Q]$  tels que, dans la fig. 2 correspondante, il n'y ait que des *carrés noirs*. [...]. [Sauvage, 1891, 312].

Si la méthode de *composition* des mineurs de Sauvage peut sembler proche de la *décomposition* des matrices de Weyr, les deux approches sont en réalité très différentes. Tandis que Sauvage s'appuie sur une analogie entre la forme du tableau et des formes géométriques, Weyr développe un processus d'itération par le calcul des puissances

*successives de compartiments d'une matrice et ce sont des opérations sur les matrices qui font émerger les mineurs et les combinent les uns par rapport aux autres.* Par opposition à la méthode statique de Sauvage, la combinatoire des matrices a un aspect *dynamique spécifique qui met la représentation en mouvement.*

A la suite des travaux de Weyr, la notion de matrice gagne une popularité croissante. La pratique polynomiale qu'elle permet sera adoptée pour la théorie des systèmes hypercomplexes par Scheffers (1890), pour la théorie des représentations de groupes par Molien (1893) et pour la théorie des formes bilinéaires par Frobenius lui-même dès 1894. On trouve des matrices dans la plupart des traités publiés en Allemagne au tournant du siècle. Pourtant, à cette époque, la représentation matricielle n'est que rarement employée et essentiellement pour représenter un déterminant. Les matrices sont notées par des symboles alphabétiques,  $A$ ,  $B$ , etc. et sont surtout employées pour représenter l'identité d'une théorie portant sur des notions distinctes (formes bilinéaires et quadratiques, substitutions linéaires, systèmes différentiels etc). Les traités d'algèbre du début du XX<sup>e</sup> siècle restent fidèles à la structure théorique de la théorie des formes bilinéaires Frobenius et aux calculs *d'invariants*. *La popularisation des matrices en Allemagne s'accompagne en fait d'une identification à la théorie des formes bilinéaires.* Cette identification se manifeste d'ailleurs, dans l'histoire présentée par les mathématiciens de l'époque, qui voient en Cayley [1858] et Frobenius [1879] *deux origines d'une même théorie* ([encart 7](#)).

La méthode matricielle élaborée par Weyr n'a, dans un premier temps, qu'une postérité très limitée et se heurte à l'existence d'une méthode paradigmatique donnant *un caractère primordial aux invariants et un rôle secondaire aux formes canoniques*, s'attachant à la tradition de Frobenius et popularisée par de nombreux traités comme ceux de Muth en Allemagne, Sauvage en France ou encore Bôcher aux Etats-Unis (<sup>39</sup>). Pour cette raison, au début du XX<sup>e</sup> siècle, *le développement de méthodes de décompositions à des formes canoniques se fait essentiellement en dehors de la théorie des formes bilinéaires et des matrices*. Développer plus avant l'histoire des matrices nécessiterait donc d'entrer dans le détail de l'histoire du théorème de « réduction des substitutions à une forme canonique », énoncé par Jordan en 1870 ([encart 13](#)) (<sup>40</sup>). Il faudrait également décrire le contexte dans lequel le "Mémoire sur le calcul fonctionnel distributif" de Pincherle de 1899 met en avant la notion *d'opération* sur des "espaces" de fonctions analytiques et pose le problème de la *décomposition d'une homographie  $A$*  comme correspondant à la recherche de *sous variétés d'un espace vectoriel invariants par rapport à  $A$*  ([encart 14](#)) (<sup>41</sup>). *Il faudrait en réalité prendre en compte de nombreux travaux ne concernant pas explicitement les matrices et attachés à des contextes différents les uns des autres.* Au tournant du XIX<sup>e</sup> siècle, *à la marge d'une théorie prédominante des invariants*, des auteurs comme Autonne, Burnside, Hensel, Jordan, Pincherle, Poincaré, Molien ou Weyr énoncent des théorèmes *distincts* au sein de réseaux *différents* et qui seront plus tard perçus comme équivalents au "théorème de Jordan de décomposition matricielle". L' [encart 15](#) présente quelques uns de ces théorèmes qui sont

<sup>39</sup> Le mémoire de Weyr lui-même fera l'objet d'une communication de Frobenius à l'Académie de Berlin en 1911 intitulée "Ueber den Rang einer Matrix" et dans laquelle Frobenius réduit la composition des matrices de Weyr à la composition des invariants polynomiaux articulés par la caractéristique de Segre.

<sup>40</sup> La postérité du théorème énoncé par Jordan en 1870 dans son *Traité des Substitutions* est tardive et passe par une ellipse : après avoir fait l'objet de nombreuses publications entre 1870 et 1874, la dénomination de "théorème de Jordan" disparaît de la scène mathématique pour réapparaître une trentaine d'années plus tard et finalement s'imposer sur un plan international dans les années 1920-1930. Dans le cadre de la théorie des formes bilinéaires en effet, la place central attribuée aux invariants implique une organisation du savoir mathématiques au sein de laquelle il n'existe pas de théorème de réduction à une forme canonique. La question du rôle respectif des invariants et des formes canoniques et d'ailleurs l'objet d'une vive querelle entre Jordan et Kronecker en 1874. Voir à ce sujet [Brechenmacher, 2006].

<sup>41</sup> La décomposition donnée par Pincherle est, d'un point de vue contemporain équivalente à celle de Weyr ; elle donne une décomposition des variétés associées aux puissances successives de chaque facteurs linéaires  $(A-z)^p$ . Contrairement aux méthodes basées sur les déterminants la méthode Weyr est valable pour les problèmes d'analyse fonctionnelle qui nécessitent le recours à la dimension infinie. Voir à ce sujet un article publié par Jean Dieudonné en 1946 : "Sur la réduction canonique des couples de matrices", *Bull. S.M.F.*, pp. 130-146.



*supportés par des pratiques de décomposition d'une représentation en tableau à des formes "canoniques", "typiques" ou "normales".*

*Décomposition, recomposition*, les méthodes distinctes du calcul des invariants et de la décomposition à des formes canoniques seront reconnues comme complémentaires dans les années 1930 et c'est en les mêlant que se construira la théorie des matrices canoniques que nous avons abordé dans la première partie. Entre 1900 et 1930, de jeunes enseignants chercheurs comme Autonne, de Séguier, Lattès ou Chatelet prêtent aux matrices des vertus pédagogiques qui leur permettent d'exposer leurs recherches les plus récentes dans des traités d'enseignement <sup>(42)</sup>. Ces préoccupations pédagogiques interrogent directement la recherche sur des questions nouvelles et impulsent le développement de la théorie des matrices canoniques par une dynamique de va-et-vient entre enseignement et recherche qui aboutira à l'élaboration d'une *culture mathématique commune* donnant un caractère *universel* à la représentation matricielle. Dans les traités des années trente, des travaux mathématiques *distincts* dans le passé sont en effet présentés comme participant d'une *même* théorie sous jacente :

Matric algebra is a mathematical abstraction underlying many seemingly diverse theories. Thus bilinear and quadratic forms, linear associative algebra (hypercomplex systems), linear homogeneous transformations and linear vector functions are various manifestations of matric algebra. Other branches of mathematics as number theory, differential and integral equations, continued fractions, projective geometry etc. make use of certain portions of this subject. Indeed , many of the fundamental properties of matrices were first discovered in the notation of a particular application, and not until much later recognized in their generality. [Mac Duffee, 1933].

Traduction. F.B.

L'algèbre des matrices est une abstraction mathématique qui sous-tend des théories diverses. Les formes bilinéaires et quadratiques, les algèbres linéaires associatives ou systèmes hypercomplexes, les transformations linéaires homogènes et les fonctions linéaires vectorielles sont ainsi différentes manifestations de l'algèbre des matrices. D'autres branches des mathématiques comme la théorie des nombres, les équations différentielles et intégrales, les fractions continues, la géométrie projective etc. font usage de ce sujet. De fait, de nombreuses propriétés fondamentales des matrices ont été d'abord découvertes au sein d'une application particulière, et seulement reconnues bien plus tard dans leur généralité.

Les méthodes matricielles des années trente articulent la *décomposition polynomiale des invariants*, un *calcul symbolique*, des *outils arithmétiques* sur les lignes et les colonnes, des *procédés d'itérations* propres à la théorie des algèbres associatives et des *méthodes vectorielles* de décomposition d'un espace en sous espaces stables. Elles réalisent une cristallisation d'idéaux et de pratiques, renvoyant à des contextes variés et qui, sur une longue durée, forgent le *caractère opératoire d'une forme de représentation*. La théorie des matrices est un élément important du processus d'unification des connaissances mathématiques qui caractérise la période 1900-1930, elle manifeste la construction d'une *culture commune à partir de pratiques auparavant distinctes*. Comme nous l'avons vu en parcourant des textes s'insérant dans un réseau donné, démêler la tresse que forme la notion contemporaine de matrice est l'occasion de présenter une *histoire plurielle de l'algèbre linéaire* qui ne se réduit pas à l'émergence de structures mais qui se veut attentive au rôle joué par *les pratiques, savoirs et idéaux propres à des réseaux, des communautés*.

---

<sup>42</sup> La période 1900-1930 ne peut être développée plus avant dans cet article. Pour une étude du rôle de Châtelet, Autonne, Lattès etc., consulter [Brechenmacher, 2006, pp. 489-652].





## ENCARTS

### **Encart 1: an introduction to the theory of canonical matrices. *Turnbul et Aitken, 1932.***

We have preferred to follow the lead of Cullis, who develops the theory in terms of the structure and properties of matrices – in **matrix idiom**, as it were, rather than in terms of **bilinear and quadratic forms, or of linear substitutions**. [...], we have, at the same time, aimed at a certain **compactness** in the formulae and demonstrations. This has been achieved in the first place by a systematic use of the matrix notation [...]. The **theory of canonical matrices** is concerned with the systematic investigation of types of transformation which **reduce** matrices to the **simplest** and most convenient **shape**. [...] in the second place, by confining the contents of each chapter almost entirely to **general theorems**, and by relegating corollaries and applications to the interspersed **sets of examples**. These examples are intended to serve not so much as exercises many being quite easy, but rather as points of relaxation, and running commentary ; they will, however, be found to contain many well-known and important theorems, which the notation establishes in the **minimum of space**. [...] The reader already familiar with the theory will also observe that certain established methods of dealing with the subject have hardly been touched upon, notably the methods of Weierstrass and Darboux, the theory of regular minors of the determinants and the treatment of quadratic forms by the method of Kronecker. We have, in fact, allowed ourselves **a free hand in dealing with the results of earlier writers, in the belief that the outcome would prove to be an easier approach** to a subject that has often failed to win affection; and the methods of H.J.S. Smith, Sylvester, Frobenius and Dickson proved in themselves quite adequate without the inclusion of other parallel theories.

[Turnbul et Aitken, 1932, 1].

#### **Plan du traité :**

##### Chapter I.

Definitions and fundamental properties of matrices.

##### Chapter II.

Elementary Transformations. Bilinear and Quadratic Forms.

##### Chapter III.

The canonical reduction of Equivalent matrices

##### Chapter IV

Subgroups of the Group of equivalent transformations.

##### Chapter V.

A Rational Canonical Form for the Collineatory group.

##### Chapter VI.

The Classical Canonical Form for the Collineatory Group.

## Encart 2 : le théorème de décomposition matricielle dans le traité de Mac Duffee de 1943.

THEOREM 65. *Every non derogatory matrix whose minimum function is  $[l(x)]^h$  where  $l(x)$  is irreducible of degree  $j$  is similar to a matrix of the type (35) with  $h$  blocks along the diagonal.*

$$(35) \quad \begin{bmatrix} \begin{array}{ccc|ccc|ccc} 0 & 0 & -l_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -l_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -l_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & -l_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -l_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -l_2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -l_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -l_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -l_2 \end{array} \end{bmatrix}$$

The case where  $F$  is the complex field, or in fact any algebraically closed field, deserves special mention. In this case we can write

$$m(x) = (x - x_1)^{h_1} (x - x_2)^{h_2} \dots (x - x_k)^{h_k}$$

so that  $A$  is similar to a direct sum of matrices of the form

$$(36) \quad B_i = \begin{bmatrix} x_i & 0 & 0 & \dots & 0 & 0 \\ 1 & x_i & 0 & \dots & 0 & 0 \\ 0 & 1 & x_i & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & x_i \end{bmatrix},$$

of  $h_i$  rows and columns. This is the familiar *Jordan normal form* of a matrix with complex elements.  
[Mac Duffee, 1943, 114-121]

### Encart 3 : extraits de la démonstration du théorème de Jordan dans le traité de Mac Duffee de 1943.

Let  $A$  and  $B$  be two square matrices,  $A$  of order  $r$  and  $B$  of order  $s$ . the matrix

$$A+B=\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

of order  $r+s$  is called their direct sum. [...] More generally, let us suppose that the  $n \times n$  matrix  $A$  of rank  $r$  can be written

$$A = A_1 + A_2 + \dots + A_k = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & A_k \end{bmatrix}$$

where  $A_i$  is of order  $n_i$  and rank  $r_i$ . Let the row **vectors** of  $A$  span the space  $S$ , let the first  $n_1$  two vectors of  $A$  span  $S_1$ , the next  $n_2$  row vectors of  $A$  span  $S_2$ , etc. [...] Thus  $S$  is the supplementary sum  $S_1 + S_2 + \dots + S_k$  of the **subspaces**  $S_i$ . [...] A subspace  $S_0$  of  $S$  is said to be an *invariant space* of the matrix  $A$  if, for every vector  $\Phi$  of  $S_0$ , it is true that  $A \cdot \Phi$  is in  $S_0$ . [...] Much of the importance of invariant spaces derive from the following result

LEMMA . Let the total vector space  $S$  be the supplementary sum of the subspaces

$$S, S, \dots, S_k$$

where each  $S_i$  is of dimension  $r_i$  and has the basis  $\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{ir_i}$ . Let  $P$  be the matrix whose column vectors are

$$(31). \sigma_{11}, \dots, \sigma_{1r_1}, \sigma_{21}, \dots, \sigma_{2r_2}, \dots, \sigma_{k1}, \dots, \sigma_{kr_k}.$$

If each space  $S_i$  is an invariant space of the matrix  $A$ , then

$$P^{-1}AP = B_1 + B_2 + \dots + B_k$$

where  $B_i$  is matrix of order  $r_i$

[...] In general  $M$  can be written

$$[\sigma_{11}, \dots, \sigma_{1r_1}, \sigma_{21}, \dots, \sigma_{2r_2}, \dots, \sigma_{k1}, \dots, \sigma_{kr_k}] \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & B_k \end{bmatrix}$$

where each submatrix  $B_i$  has  $r_i$  rows and columns.

[...]

THEOREM 64. Let  $A$  be any  $n \times n$  matrix with elements in a field  $F$ , and let

$$m(x) = m_1(x)m_2(x)\dots m_k(x)$$

be its minimum function expressed as a product of polynomials which are relatively prime as pairs. Let the null space of  $m_i(A)$  be of rank  $r_i$ . then  $A$  is similar to a direct sum

$$B_1 + B_2 + \dots + B_k$$

where  $B_i$  is of order  $r_i$ , and the minimum function of  $B_i$  is  $m_i(x)$ .

[...] Example :

$$A = \begin{bmatrix} -1 & 5 & 5 & 3 \\ 5 & 0 & -17 & -26 \\ -10 & -2 & 32 & 51 \\ 6 & 2 & -19 & -31 \end{bmatrix}$$

Let us choose  $\Phi = (1, 0, 0, 0)$ . Then

$$P = (\Phi, A\Phi, A^2\Phi, A^3\Phi) = \begin{bmatrix} 1 & -1 & -6 & 5 \\ 0 & 5 & 5 & 0 \\ 0 & -10 & -10 & 2 \\ 0 & 6 & 6 & 0 \end{bmatrix}$$

[...] We find that

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

which is the companion matrix of the minimum equation

$$m(x) = x^4 + x^2 + 1 = 0$$

The second canonical form can be obtained from the matrix

$$B_1 = \begin{bmatrix} \frac{2}{3} & \frac{19}{3} \\ \frac{3}{1} & -\frac{5}{3} \\ -\frac{1}{3} & -\frac{5}{3} \end{bmatrix}, B_2 = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{2}{7} & -\frac{1}{2} \\ -\frac{7}{2} & -\frac{1}{2} \end{bmatrix}$$

Let

$$Q_1 = \begin{bmatrix} 1 & \frac{2}{3} \\ 0 & -\frac{1}{3} \end{bmatrix}, Q_2 = \begin{bmatrix} 1 & \frac{3}{2} \\ 0 & -\frac{7}{2} \end{bmatrix}.$$

Then

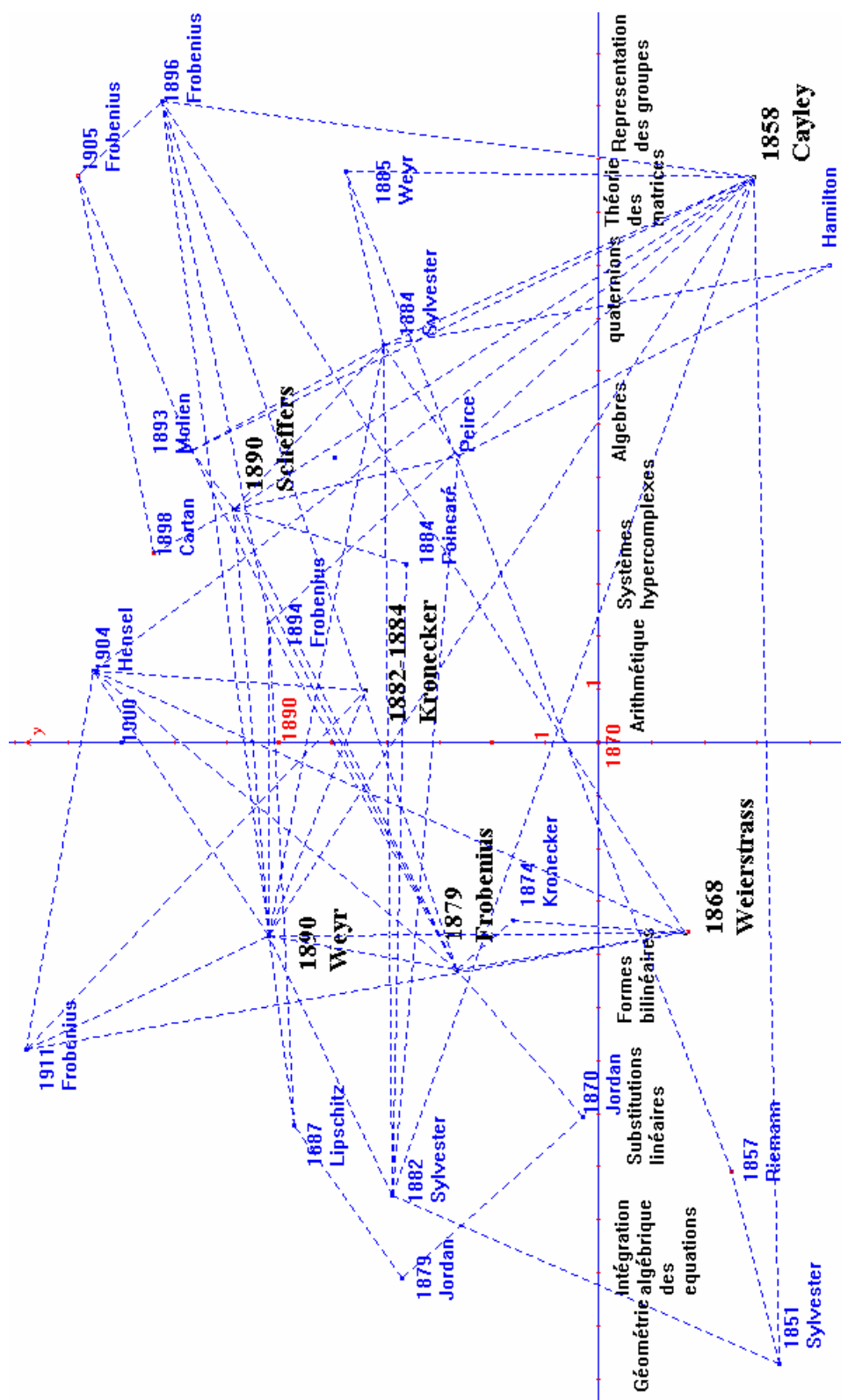
$$Q_1^{-1}B_1Q_1 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, Q_2^{-1}B_2Q_2 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

are the companion matrices of the equations obtained from the respective factors  $x^2+x+1$  and  $x^2-x+1$  of the minimum function of  $A$ . Then  $A$  is similar to the matrix

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

**Encart 4 : un exemple de représentation de réseaux de recherches (1880-1907).**

Dans l'exemple de réseau représenté ci-dessous, le *Traité des substitutions* de Jordan joue un rôle marginal quand on le situe par rapport aux nœuds que sont les mémoires de Frobenius [1879], Weierstrass [1858] et Weyr [1890] et sur lesquels nous nous attarderons dans la deuxième partie de cet article. On peut remarquer qu'à partir de 1890, Cayley [1858] devient un nœud *a posteriori* et se trouve relié à Frobenius [1879].





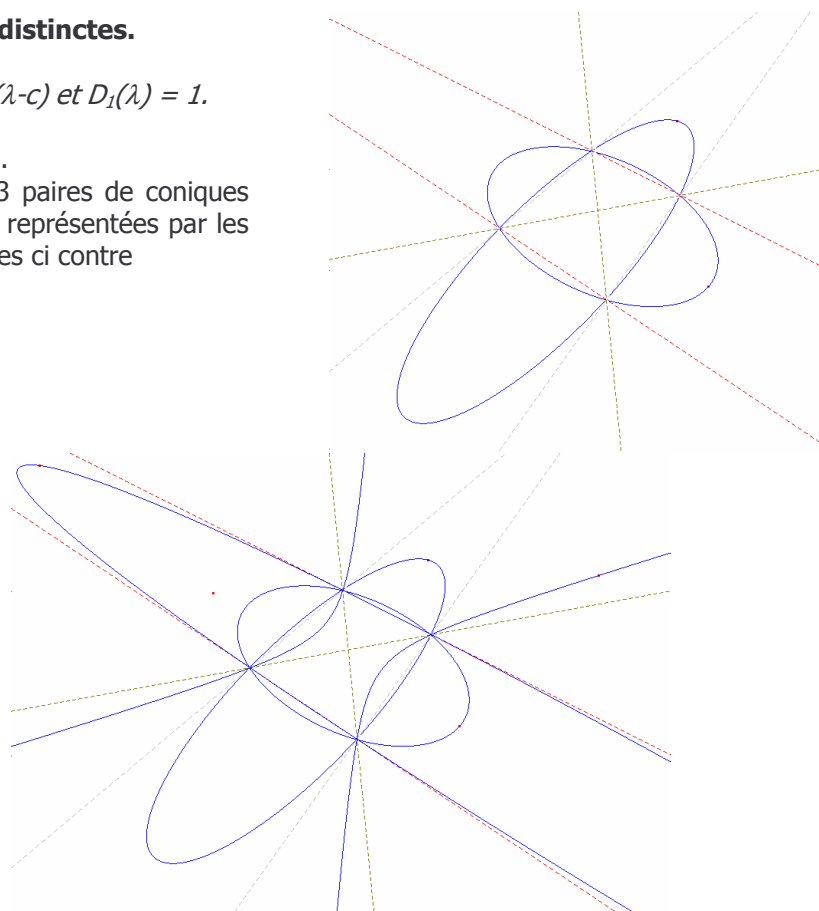
## Encart 5 : les types d'intersections des coniques et la décomposition polynomiale du déterminant.

### 1<sup>er</sup> type : 3 racines distinctes.

$$D(\lambda) = (\lambda-a)(\lambda-b)(\lambda-c) \text{ et } D_I(\lambda) = 1.$$

3 points d'intersection.

Le faisceau contient 3 paires de coniques dégénérées distinctes représentées par les trois couples de droites ci contre



Sylvester caractérise cette situation de la façon suivante :

When there is no contact, we have as the types of their characteristics

$$U = x^2 + y^2 + z^2$$

$$V = ax^2 + by^2 + cz^2$$

The three roots of  $(V + \lambda U) = 0$  are  $\lambda = -a$ ,  $\lambda = -b$ ,  $\lambda = -c$ ,

Showing that there are three distinct pairs of lines in which the intersections of  $U$  and  $V$  are contained, the equations to three pairs being respectively :

$$(b-a)y^2 + (c-a)z^2 = 0,$$

$$(c-b)z^2 + (a-b)x^2 = 0,$$

$$(a-c)x^2 + (b-c)y^2 = 0 ;$$

the four points of the intersection being defined by the equations corresponding to the proportions

$$x : y : z :: \sqrt{b-c} : \sqrt{c-a} : \sqrt{a-b} \gg$$

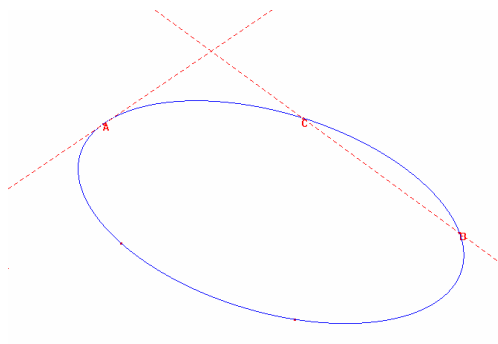
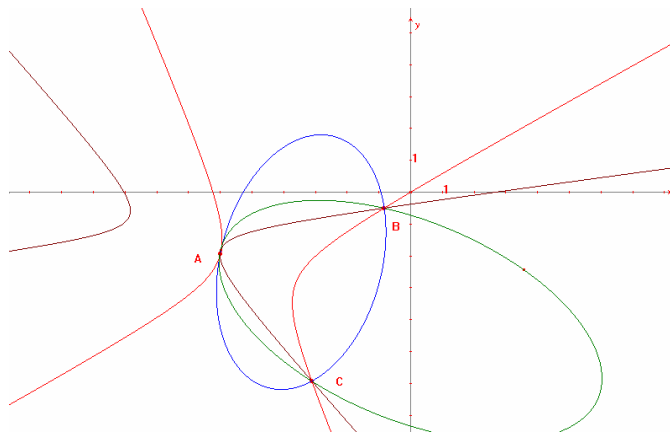
## 2<sup>e</sup> type (simple contact) : une racine double.

Déterminant et mineurs sans facteurs communs.

$$D(\lambda) = (\lambda-a)^2(\lambda-b) \text{ et } D_1(\lambda) = 1$$

Un point d'intersection double et 2 points simples.

Le faisceau contient 2 paire de coniques dégénérées distinctes : la tangente D et (BC) ainsi que (AB) ( $U+\lambda V$  a deux valeurs propres distinctes).



Sylvester caractérise cette situation de la façon suivante :

Now let  $(U+\lambda V)$  have two equal roots ; the characteristics assume the form

$$U = x^2 + y^2 + xz,$$

$$V = ax^2 + by^2 + cxz$$

Two of the pairs of lines become identical, that is two of the four points of intersection coincide.

This may be termed 'Simple contact ». The tangent at the point of contact is  $x=0$  ; this equation making  $U$  and  $V$  each become of only one order.

The intersections are

$$x=0, y=0,$$

$$x=0 y=0$$

$$\sqrt{(a-c)x} + \sqrt{(b-c)y} = 0, z=0,$$

$$\sqrt{(a-c)x} - \sqrt{(b-c)y} = 0, z=0.$$

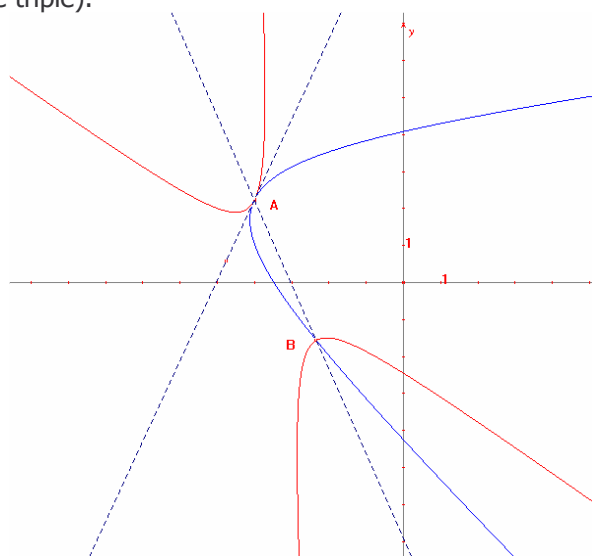
### 3<sup>e</sup> type (proximal contact) : trois racines doubles.

Déterminant et mineurs sans facteurs communs.

$$D(\lambda) = (\lambda - a)^3 \text{ et } D_I(\lambda) = 1$$

1 point d'intersection triple.

Le faisceau contient une paire de coniques dégénérées (la tangente et la droite (AB),  $U + \lambda V$  a une valeur propre triple).



Now let  $(U + \lambda V)$  have all its roots equal. This condition will be satisfied [...] by making

$$U = x^2 + yz + yx,$$

$$V = ax^2 + ayz + byx$$

Here only one distinct pair of lines can be drawn to contain the intersections, showing that three out of the four points come together.

This may be termed « Proximal Contact. » The number of affirmative conditions to be satisfied is two, and the contact is therefore entitled of the second degree.

The tangent at the point of contact is  $y=0$ , and the four intersections become

$$x=0, y=0,$$

$$x=0, y=0,$$

$$x=0, y=0,$$

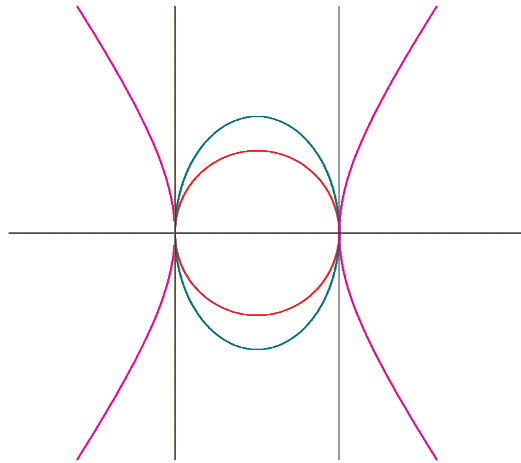
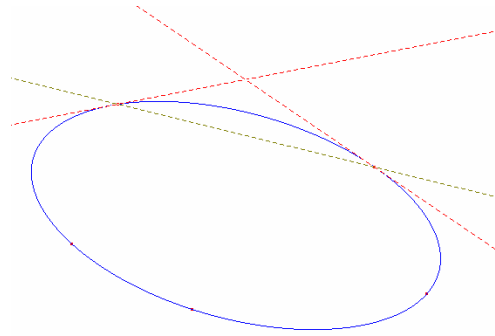
$$x=0, z=0.$$

#### 4<sup>e</sup> type (diploidal contact) : une racine double.

Déterminant et mineurs avec facteurs communs.

$$D(\lambda) = (\lambda-a)^2(\lambda-b) \text{ et } D_1(\lambda) = \lambda-a$$

2 points d'intersections doubles. Le faisceau contient 2 paire de coniques dégénérées distinctes : les deux tangentes et une droite double. ( $U+\lambda V$  a deux valeurs propres distinctes).



Let now the first minors of  $U+\lambda V$  have a factor in common ; this will indicate that  $U+\lambda V$  may be made to lose two orders by rightly assigning  $\lambda$ , in other words, that the intersections of  $U$  and  $V$  be contained upon a pair of coincident lines. Here it is remarkable that the original forms of  $U$  and  $V$  reappear, but with a special relation of equality between the coefficients : we shall have, in fact,

$$U = x^2 + y^2 + z^2, \\ V = ax^2 + ay^2 + bz^2$$

This gives the law for double, or, as I prefer to call it, diploidal contact. By virtue of the Homaloidal law, we know that if three first minors of  $U+\lambda V$  be zero, all are zero ; we have therefore to express that three quadratic functions of  $\lambda$  have a root in common. This implies the existence of two affirmative conditions; the contact of the two conics taken collectively may therefore be still entitled of the second degree, although the contact at each of the two points where it takes place is simple, or of the first degree.

These points are evidently defined by the equation

$$\begin{aligned} \{x + \sqrt{-1}y = 0, z=0\}, \\ \{x - \sqrt{-1}y = 0, z=0\}, \end{aligned}$$

and the ordinary algebraical solutions of the equations  $U=0, V=0$  would naturally lead to the four systems

$$\begin{aligned} x + \sqrt{-1}y = 0, z=0 \\ x + \sqrt{-1}y = 0, z=0 \\ x - \sqrt{-1}y = 0, z=0, \\ x - \sqrt{-1}y = 0, z=0 \end{aligned}$$

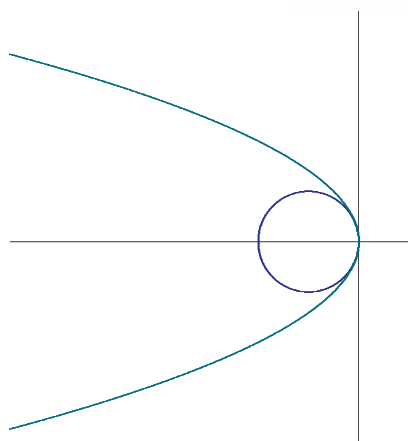
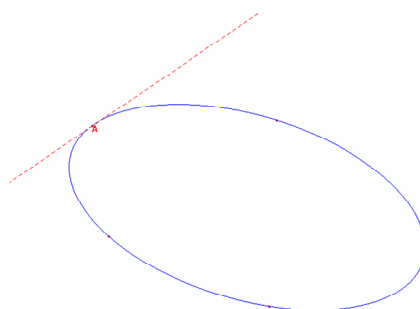
The two tangents at the point of contact are  $x + \sqrt{-1}y = 0, x - \sqrt{-1}y = 0$  and the coincident pair of lines containing the intersections is  $z^2 = 0$ .

### 5<sup>e</sup> type (confluent contact) : 1 racine triple.

Déterminant et mineurs avec facteur commun.

$$D(\lambda) = (\lambda - a)^3 \text{ et } D_1(\lambda) = (\lambda - a)$$

1 point d'intersection quadruple. Le faisceau contient 1 paire double de coniques dégénérées : la tangente en A ( $U + \lambda V$  a une valeur propre).



In the ordinary case of diploidal contact, the first minors having a factor in common, this factor will enter twice into the complete determinant of  $U + \lambda V$ , but it may enter three times : this will indicate, that not only do the four intersections lie on a coincident pair of lines, but furthermore, that there is but one pair of lines of any kind on which they lie.

In the ordinary case of diploidal contact, it will be observed that this latter condition does not obtain ; the four intersections lie on a coincident pair of lines ; but they lie also on a crossing pair, namely, in two tangents at the point of contact. In this higher species of diploidal contact, it is clear that the two points of contact, which are ordinarily distinct, come together, and that all four intersections coincide.

This I call confluent contact ; the forms  $U$  and  $V$  corresponding thereto will be

$$U = x^2 + y^2 + xz,$$

$$V = ay^2 + axz$$

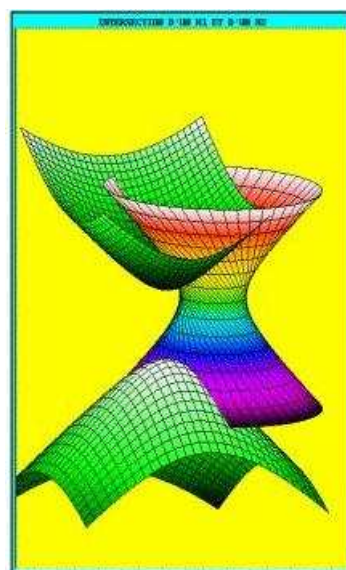
The common tangent at the point of contact being  $x=0$ , and the four coincident points  $x^2 = 0, y^2 = 0$ .

The number of affirmative conditions to be satisfied being three, the contact is to be entitled of the third degree.

Observe, that it is of no use to descend below the first minors in this case ; because the second minors, being linear functions of  $\lambda$ , could not have a factor in common, unless  $V:U$  becomes a numerical ratio, which would imply that the conics coincided.

## Encart 6 : les types d'intersection des quadriques

- **Les cas regroupés dans la « 1<sup>st</sup> Class »**  
correspondent à  $D_1(\lambda) = I$  :
  - simple contact (4 racines distinctes),
  - proximal contact (1 racine double),
  - diploïdal contact (1 racine triple),
  - confluent contact (1 racine quadruple).



- **Les cas regroupés dans la « 2<sup>nd</sup> Class »**  
correspondent à un facteur commun linéaire pour les premiers mineurs :  $D_1(\lambda) \neq I$  mais  $D_2(\lambda) = I$  :

« 1<sup>st</sup> Section »  $D_1(\lambda) = \lambda - a$

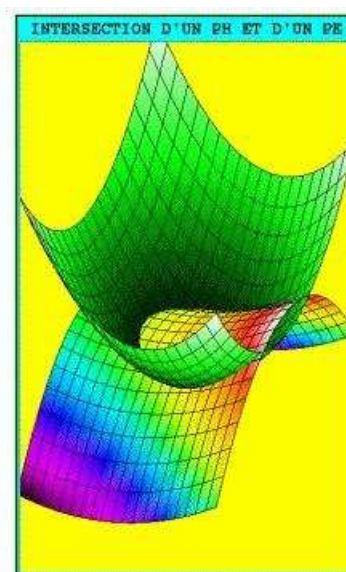
- biplanar contact :  $D_1(\lambda) = \lambda - a$  : et  $D(\lambda) = (\lambda - a)^2(\lambda - b)\lambda - c$ ,
- confluent biplanar contact :  $D_1(\lambda) = \lambda - a$  et  $D(\lambda) = (\lambda - a)^3(\lambda - c)$ ,
- bilinear contact :  $D_1(\lambda) = \lambda - a$  et  $D(\lambda) = (\lambda - a)^2(\lambda - b)^2$ ,
- confluent-bilinear contact :  $D_1(\lambda) = \lambda - a$  et  $D(\lambda) = (\lambda - a)^4$ ,

« 2<sup>nd</sup> section »  $D_1(\lambda) = (\lambda - a)(\lambda - b)$

- quadrilinear, or quadruple contact :  $D(\lambda) = (\lambda - a)^2(\lambda - b)^2$  et  $D_1(\lambda) = (\lambda - a)(\lambda - b)$ ,
- unilineo-indefinite contact :  $D(\lambda) = (\lambda - a)^4$  et  $D_1(\lambda) = (\lambda - a)^2$ ,

« 3<sup>rd</sup> Class » :  $D_2(\lambda) = \lambda - a$  et  $D_1(\lambda) = (\lambda - a)^2$

- curvilinear-indefinite contact.  $D_2(\lambda) = \lambda - a$ ,  $D_1(\lambda) = (\lambda - a)^2$  et  $D(\lambda) = (\lambda - a)^3(\lambda - b)$ ,
- bilinear-indefinite contact.  $D_2(\lambda) = \lambda - a$ ,  $D_1(\lambda) = (\lambda - a)^2$  et  $D(\lambda) = (\lambda - a)^4$ .





## **Encart 7 : les perceptions du mémoire de Cayley (1858) à la fin du XIX<sup>e</sup> siècle et l'écriture d'une histoire.**

Dans les années 1880-1890, une histoire de la théorie des matrices est écrite en direct par les mathématiciens et s'étoffe à mesure que la notion de matrice évolue. En quelques années, Sylvester abandonne sa conception originelle d'une matrice comme mère des mineurs d'un déterminant pour définir une matrice comme une quantité complexe satisfaisant une certaine équation algébrique. La théorie de Sylvester d'une "algèbre universelle" englobant les substitutions linéaires, les homographies de Babbage et les quaternions de Hamilton fait participer d'une même théorie des notions jusqu'alors distinctes et implique un effet rétroactif sur l'histoire des mathématiques. La dénomination de "théorème de Cayley-Hamilton" illustre la façon dont l'adoption par Sylvester d'une conception des matrices comme quantité complexe vérifiant une équation algébrique, s'accompagne de la reconnaissance d'une priorité double sur l'énoncé d'un théorème :

C'est dans les *Lectures*, publiées en 1844 que pour la première fois a paru la belle conception de l'équation identique appliquée aux matrices du troisième ordre, enveloppée dans un langage propre à Hamilton, après lui mise à nu par M. Cayley dans un très important Mémoire sur les matrices dans les *Philosophical Transactions* pour 1857 ou 1858, et étendue par lui aux matrices d'un ordre quelconque, mais sans démonstration ; [...]. [Sylvester, 1884j, 202].

L'évolution de la notion de matrice de Sylvester entre 1882-1885 entraîne une évolution parallèle du rôle historique du mémoire de Cayley. La première lecture de Cayley par Sylvester est due à la postérité d'un mathématicien lu pour sa méthode de résolution d'un problème précis : le calcul des racines des substitutions. En 1884, Sylvester réévalue le rôle historique de Cayley désormais considéré comme précurseur de l'"algèbre universelle", science de la "quantité multiple". C'est à ce moment que le rôle historique de Cayley n'est plus perçu comme attaché à l'énoncé d'une méthode de calcul des racines de substitutions mais à la définition des opérations :

I had seen the memoir [...] and indeed forced itself to my attention as a means of giving simplicity and generality to my formula for the powers or roots of matrices (1882) [...] I was led to the discovery of the properties of the latent roots of matrices and had made considerable progress in developing the theory of matrices considered as quantities, when on writing to Prof Cayley upon the subject he referred me to the memoir in question, all this only proves how far the discovery of the quantitative nature of matrices is removed from being artificial or factitious, but on the contrary, was bound to be evolved, in the fullness of time, as a necessary sequel to previously acquired cognitions . [Sylvester, 1885].

Les premières lignes de l'histoire qui voit dans le calcul des matrices de Cayley une origine de l'étude théorique des algèbres associatives sont écrites par Sylvester lui-même. Eduard Weyr y ajoute un chapitre lorsque, dans l'introduction de son mémoire de 1890, il fait participer les matrices de Cayley et les formes bilinéaires de Frobenius d'une même théorie et, du même coup, partage les priorités. Quelques exemples : la "seconde loi du mouvement algébrique" de Sylvester qui a suscité les premiers travaux de Weyr sur les matrices en 1884-1885 est attribuée à Frobenius; l'équation aux racines latentes de Sylvester est identifiée à l'équation caractéristique de Cauchy. Par une réorganisation du savoir mathématique du moment, Weyr réorganise aussi l'histoire. La manifestation la plus exemplaire en est qu'à partir du mémoire de Weyr, Frobenius est considéré comme ayant donné la première démonstration générale du théorème de Cayley Hamilton. Le développement de la théorie des systèmes hypercomplexes dans le monde anglo-saxon ([Taber, 1889], [Bucheim, 1885])

comme sur le continent ([Study, 1889], [Scheffers, 1891], [Molien, 1893] et [Frobenius, 1894]), donne à l'histoire des matrices sa structure quasi-définitive. Aux références à Cayley et Hamilton se trouvent associés les travaux des Peirce et de Clifford sur les algèbres associatives ainsi que les travaux géométriques de Grassmann (<sup>43</sup>).

Accordingly, Cayley laid down the laws of combination of matrices upon the basis of the combined effect of the matrices as operators of linear transformation upon a set of scalar variables or carriers. The development of the theory, as contained in Cayley's memoir, was the development of the consequences of these primary laws of combination. Before Cayley's memoir appeared, Hamilton had investigated the theory of such a symbol of operation as would convert three vectors into three linear functions of those vectors, which he called a linear vector operator. Such an operator is essentially identical with a matrix as defined by Cayley; and some of the chief points in the theory of matrices were made out by Hamilton and published in his *Lectures on Quaternions* (1852). [...] Hamilton must be regarded as the originator of the theory of matrices, as he was the first to show that the symbol of a linear transformation might be made the subject-matter of a calculus. [...] the identity of the two theories was explicitly mentioned by Clifford in a passage of his *Dynamic*, and was virtually recognized elsewhere by himself and by Tait. Sylvester carried the investigation much farther, developing the subject on the same basis as that which Cayley had adopted. Subsequent to Cayley, but previous to Sylvester, the Peirces, especially Charles Peirce, were led to the consideration of matrices from a different point of view ; namely, from the investigation of linear associative algebra involving any number of linearly independent units. [...] In the present paper I regard a matrix as a linear unit operator, operating upon the linearly independent units of an algebra, without reference to any meaning of such units, or to any properties which these units may have in combination with each other ; and I have in this way endeavoured to develop the theory of matrices. Taber [1889, 337].

La théorie des systèmes de nombres complexes a son origine dans les travaux de Gauss de 1831 qui provoquèrent l'admission définitive des nombres complexes dans la science. [...] Aussi Hamilton est d'abord allé vers une interprétation géométrique [des nombres complexes] comme points de l'espace et est arrivé en 1843 à ses quaternions [...]. Le premier à introduire une étude d'ampleur générale des systèmes de nombres, est H. Grassmann dans son *Ausdehnungslehre* de 1862. [...] Les développements anglo-américains, comme ceux de Cayley, Clifford, Peirce et Sylvester [...] montrent l'importance de l'associativité de la loi de multiplication [...], de plus ils considèrent la multiplication comme une substitution linéaire. Un traitement semblable des nombres complexes et des substitutions linéaires est alors donné avec succès par Frobenius en 1878. [Scheffers, 1891, 386-388, traduction, F.B.].

Le rôle de précurseur des algèbres associatives attribué à Cayley est indissociable de l'organisation de la nouvelle théorie qui présente les matrices comme cas particulier de la théorie des systèmes hypercomplexes.

---

<sup>43</sup> Pour la théorie des grandeurs complexes de Sylvester, l'influence principale semble provenir des Peirce. Benjamin Peirce (1809-1880) publie en 1870, dans la tradition de la théorie des quaternions qu'il enseigne à Harvard, une centaine de copies de son traité, intitulé "Linear associative Algebra". Peirce détermine 162 algèbres associatives sur le corps complexe, de dimension inférieure ou égale à 6, et introduit les notions d'éléments idempotents et nilpotents. Charles Sander Peirce (1839-1914) réalise des travaux de logique influencés par les algèbres associatives et annote les travaux de son père avant de les publier dans le journal de Sylvester en 1881. Le mémoire de Cayley n'exerce cependant aucune influence sur les travaux des Peirce [Parshall, 1985].

## Encart 8 : Quelques éléments biographiques sur Eduard Weyr.

- **Sur la carrière de Weyr.**

- Dissertation donnée à Göttingen en 1873 sur le thème "Ueber algebraische Raumkurven".
- Débute sa carrière en suivant les traces de son frère aîné, Emil Weyr, qui travaille essentiellement sur la géométrie et notamment sur les involutions. Emil inaugure la première séance de la *S.M.F* en 1870 [Gispert, 1991, 20], devient membre du conseil non résident et fait don de ses mémoires mathématiques à la société en 1873. Eduard est membre de la société à partir de 1873.
- 1873-74. Etudiant à l'université de Karl Ferdinand de Prague (*Bulletin de la Société Mathématique de France*, 2).
- 1878. Professeur à l'université de Prague (*Bulletin de la Société Mathématique de France*, 6 ([1878](#)), 13).
- 1880-1904. Professeur à l'école polytechnique de Prague à partir de 1880 (*Bulletin de la Société Mathématique de France*, [23](#) ([1895](#)), 13).
- A partir de 1878, rédacteur de la société savante *Casopis pro matematiky a fysiky*, cette société publie un journal dans lequel Weyr intervient le plus souvent sur des questions de géométrie ou d'analyse.

- **Travaux mathématiques et implication internationale de Weyr.**

Dans les années 1880, Eduard Weyr publie des mémoires de géométrie à Paris et à Vienne (plans tangents et courbes osculatrices). Ces travaux l'amènent à traiter dans les *Casopis* des questions de calcul des déterminants [Muir, 1906, 5 et 100]. A partir de 1884, Weyr débute des travaux sur les quaternions à la suite des récentes communications de Sylvester à l'Académie de Paris. Les premières notes dans les *Comptes Rendus* sont développées par des mémoires sur les systèmes hypercomplexes publiés dans le *Bulletin des Sciences Mathématiques* et à l'Académie de Prague [Weyr, 1887a,b,c]. Les travaux de la période 1884-1890 sont réorganisés sous forme synthétique par la mémoire de 1890 sur les formes bilinéaires dont il est question dans cette partie.

Weyr participe au congrès international des mathématiciens de Chicago de 1893 avec un exposé "Sur l'équation des lignes géodésiques".

Weyr est le correspondant de Hermite à Prague. Les deux savants ont des échanges portant sur les fonctions de variables réelles (voir la correspondance de Hermite - Œuvres II 393-400, 496-499) qui aboutissent à l'énoncé d'un théorème dit de Hermite-Weyr.

## Encart 9 : plan d'une démonstration contemporaine du théorème de Jordan.

La démonstration contemporaine la plus répandue de la décomposition de Jordan d'un endomorphisme  $f \in \text{End}_K(E)$  de polynôme caractéristique scindé découle d'une décomposition d'un espace vectoriel  $E$  de dimension finie sur un corps  $K$  en sous espaces caractéristique sous l'action de l'opérateur  $f$ .

- Théorème de décomposition des noyaux.**

Soit  $F(x)$  un polynôme tel que  $F(f)=0$  ; on suppose que  $F(X)=F_1(X)\dots F_k(X)$  où les  $F_i$  sont 2 à 2 premiers entre eux. On pose  $E_i=\text{Ker}F_i(f)$ . Alors les sous espaces  $E_i$  sont stables par  $f$ , on a  $E=E_1+E_2+\dots+E_k$  et les projecteurs  $p_i : E \leftrightarrow E$  tels que  $\text{Imp}_i = E_i$  sont dans  $K[f]$ , c'est-à-dire polynomiaux en  $f$ .

- Application aux polynômes minimal et caractéristique.**

Soit  $\chi_f(X)$  le polynôme caractéristique de  $f$  et  $M(X)$  le polynôme minimal.

$$\chi_f(X) = \prod_{1 \leq i \leq k} (X - \lambda_i)^{n_i}, M(X) = \prod_{1 \leq i \leq k} (X - \lambda_i)^{m_i}$$

où  $m_i \hat{=} n_i$ . Alors  $E_i = \text{Ker}(f - \lambda_i I)^{n_i} = \text{Ker}(f - \lambda_i I)^{m_i}$  et  $(X - \lambda_i I)^{m_i}$  est le polynôme minimal de l'endomorphisme  $f_i$  induit par  $f$ .

$$E = E_1 + E_2 + \dots + E_k.$$

Pour que  $f$  soit diagonalisable, il faut et il suffit que les valeurs propres  $\lambda_i$  soient racines simples du polynôme minimal  $M(X)$ .

- Endomorphismes nilpotents.**

On suppose  $f^m \neq 0$  et  $f^{m+1} = 0$ . Soit  $E_i = \text{Ker}f^i = \{x/f^i(x)=0\}$ . On a la suite d'inclusions strictes :

$$E_0 = \{0\} \subset E_1 \subset E_2 \subset \dots \subset E_{m+1} = E.$$

Pour tout  $i$ ,  $f(E_{i+1}) \subset E_i$  et  $F \cap E_{i+1} = \{0\} \iff f(F) \cap E_i = \{0\}$ .

Soit  $F_m$  le supplémentaire de  $E_m$  dans  $E_{m+1}=E : E = E_{m+1} = F_m + E_m$  et par récurrence, ayant  $F_i$  tel que :  $E_{i+1} = F_i + E_i$ ,  $f(F_i) \subset E_i$ , on a  $f(F_i) \cap E_{i+1} = \{0\}$ . On forme  $F_{i-1}$  le supplémentaire de  $E_{i-1}$  dans  $E_i$  contenant  $f(F_i)$ .

$$E = F_0 + F_1 + \dots + F_m$$

et  $f$  induit une injection  $F_i \hookrightarrow F_{i-1}$ ,  $1 \hat{=} i \hat{=} m$ .

Soit  $e_n$ , un vecteur non nul de  $F_m$ , on forme les vecteurs

$$e_{n-1} = f(e_n), \dots, e_{n-m} = f^m(e_n).$$

Si  $\dim F_m > 1$ , on prend  $e_{n-m-1}$ , non colinéaire à  $e_n$  et on recommence.

Une fois les vecteurs de  $F_m$  épuisés, si  $f(F_m)$  est strictement contenu dans  $F_{m-1}$ , on continue en prenant un vecteur dans  $F_{m-1} - f(F_m)$  puis ses images successives. On recommence le processus jusqu'à épuisement de tous les vecteurs de  $F_0$  et on obtient ainsi une base  $e$  de  $E$  dans laquelle la matrice a la forme suivante :  $M = (x_{ij})$  où  $x_{ij} = 0$  sauf pour les termes de la forme :  $y_i = x_{i,i+1}$  qui valent 1 ou 0.

- Forme matricielle de Jordan.**

Il existe une base de  $E$  par rapport à laquelle la matrice de  $f$  soit diagonale par blocs, chaque bloc étant somme d'une matrice scalaire et d'une matrice du type précédent. Cette réduction est la forme réduite de Jordan de l'endomorphisme  $f$ .

Exemple :  $\dim F_3 = 1, \dim F_2 = 2, \dim F_1 = 3, \dim F_0 = 3$ .

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

## Encart 10: le rôle de la représentation matricielle dans la formulation contemporaine de la relation mathématique entre diviseurs élémentaires et forme de Jordan.

Trois exemples de décompositions matricielles associées à la décomposition en diviseurs élémentaires d'un même polynôme caractéristique :  $|A-\lambda I| = (\lambda-1)^2(\lambda-2)^3(\lambda-3)$ .

Forme de Jordan	$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 & \\ & & & & & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & 1 & \\ & & & 2 & 1 \\ & & & & 2 \\ & & & & & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & 0 & \\ & & & 2 & 1 \\ & & & & 2 \\ & & & & & 3 \end{pmatrix}$
Diviseurs élémentaire	$(\lambda-1), (\lambda-1),$ $(\lambda-2), (\lambda-2), (\lambda-2),$ $(\lambda-3)$	$(\lambda-1), (\lambda-1),$ $(\lambda-2)^3,$ $(\lambda-3)$	$(\lambda-1), (\lambda-1),$ $(\lambda-2), (\lambda-2)^2,$ $(\lambda-3)$

### Extraits du traité de Wedderburn de 1933.

[Wedderburn, 1933] :

Theorem 6. If  $\lambda_1, \lambda_2, \dots, \lambda_s$  are any constants, not necessarily all different and  $v_1, v_2, \dots, v_s$  are positive integers whose sum is  $n$ , and if  $a_i$  is the array of  $v_i$  rows and columns given by

$$\begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \end{pmatrix} \quad (10)$$

where each coordinate on the main diagonal equals  $\lambda_i$ , those on the parallel on its right are 1, and the remaining ones are 0, and if  $a$  is the matrix of  $n$  rows and columns given by

$$a = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \dots & \\ & & & a_s \end{pmatrix} \quad (11)$$

composed of blocks of terms defined by (10) arranged so that the main diagonal of each lies on the main diagonal of  $a$ , the other coordinates being 0, then  $\lambda-a$  has the elementary divisors

$$(\lambda - \lambda_1)^{v_1}, (\lambda - \lambda_2)^{v_2}, \dots, (\lambda - \lambda_s)^{v_s} \quad (12)$$

[...]

If  $A$  is a matrix with the same elementary divisors as  $a$ , it follows from theorem 5, that there is a matrix  $P$  such that  $A = PaP^{-1}$  and hence, if we choose in place of the fundamental basis  $(e_1, e_2, \dots, e_n)$  the basis  $(Pe_1, Pe_2, \dots, Pe_n)$ , [...] (11) gives the form of  $A$  relative to the new basis. This form is called the canonical form of  $A$ .

## Encart 11: comparaison du calcul des matrices de Weyr et de la théorie des matrices de Cayley.

### *Structure du mémoire de Cayley*

Définition des matrices.  
Opérations sur les matrices.  
Enoncé du "théorème remarquable" : "une matrice satisfait une équation algébrique de son propre ordre".  
Equations en matrices.  
Matrices périodiques.  
Matrices permutables.  
Matrices alternées et symétriques  
Puissances de matrices.

### *Structure de la première partie du mémoire de Weyr*

1. Le calcul avec les matrices
2. La composition des systèmes de quantités.
3. Sur la nullité des matrices.
4. Les racines d'une matrice et les nombres caractéristiques.
5. L'équation fondamentale d'une matrices.

La seconde partie du mémoire de Weyr ne présente aucune référence à celui de Cayley et concerne des questions étrangères à la théorie de Cayley comme la transformation des formes bilinéaires et son application aux équations différentielles linéaires (numéros 7, 9, 10, 11).



## Encart 12 : comparaisons entre Weyr et Cayley

### Comparaison des définitions des matrices chez Weyr et Cayley.

La définition des matrices par Weyr.	La définition des matrices par Cayley.
<p>Eine Matrix <math>n^{ter}</math> Ordnung <math>\{a_{hk}\}</math> oder <math>A</math> auf ein System von <math>n</math> Werten <math>x_1, \dots, x_n</math> oder kürzer <math>(x)</math> applicieren, heißt ein Werthsystem <math>y_1, \dots, y_n</math> mittelst der <math>n</math> linearen Gleichungen</p> $y_h = a_{h1}x_1 + \dots + a_{hn}x_n \quad (h=1, 2, \dots, n)$ <p>ableiten. Diese <math>n</math> Gleichungen mögen durch die symbolische Gleichung</p> $(y) = A(x)$ <p>ausgedrückt werden.</p>	<p>The term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, [...]; in this restricted sense, a set of quantities arranged in the form of a square, e.g.</p> $\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$ <p>is said to be a matrix. [...]</p> <p>The notation</p> $\begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} (x, y, z),$ <p>represents the set of linear functions</p> $((a, b, c)(x, y, z), (a', b', c')(x, y, z), (a'', b'', c'')(x, y, z),$ <p>so that calling these <math>(X, Y, Z)</math>, we have</p> $(X, Y, Z) = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} (x, y, z),$

## Comparaison du calcul des matrices de Weyr et des opérations sur les matrices de Cayley.

Calcul des matrices de Weyr.	Opérations sur les matrices de Cayley.
<p>Als summe zweier Matrizen <math>n^{ter}</math> Ordnung  <math>A = \{a_{hk}\}, A' = \{a'_{hk}\}</math>  wird jene Matrize <math>B = \{b_{hk}\}</math> bezeichnet, die auf ein beliebiges System <math>(x)</math> appliciert das System <math>(y+y')</math> ergibt, wobei <math>(y)</math> und <math>(y')</math> die aus <math>(x)</math> durch <math>A</math> und <math>A'</math> abgeleiteten Systeme bezeichnen.  Hieraus folgt sofort  <math display="block">b_{hk} = a_{hk} + a'_{hk},</math> d.h. die Elemente der Summe sind Summen der correspondierenden Elemente der Summanden.  Eine Matrix, deren Elemente durchwegs gleich Null sind, möge Nullmatrix heißen und mit <math>O</math> bezeichnet werden.  Definiert man analog die Differenz zweier Matrizen, so findet man  <math display="block">A - A' = B</math> falls  <math display="block">a_{hk} - a'_{hk} = b_{hk} \quad (h, k = 1, 2, \dots, n).</math> Offenbar gilt  <math display="block">A - A' + A' = A, \quad A - A = O.</math> Dans Symbol <math>-A</math> möge die Differenz <math>O - A</math> bedeuten; es hat demnach <math>-A</math> die Elemente <math>-a_{hk}</math></p>	<p>The equations  <math display="block">(X, Y, Z) = \begin{vmatrix} a &amp; b &amp; c \\ a' &amp; b' &amp; c' \\ a'' &amp; b'' &amp; c'' \end{vmatrix} (x, y, z),</math> <math display="block">(X', Y', Z') = \begin{vmatrix} \alpha &amp; \beta &amp; \gamma \\ \alpha' &amp; \beta' &amp; \gamma' \\ \alpha'' &amp; \beta'' &amp; \gamma'' \end{vmatrix} (x, y, z),</math> give  <math display="block">(X + X', Y + Y', Z + Z') = \begin{vmatrix} a + \alpha &amp; b + \beta &amp; c + \gamma \\ a' + \alpha' &amp; b' + \beta' &amp; c' + \gamma' \\ a'' + \alpha'' &amp; b'' + \beta'' &amp; c'' + \gamma'' \end{vmatrix} (x, y, z),</math> and this lead to  <math display="block">\begin{vmatrix} a + \alpha &amp; b + \beta &amp; c + \gamma \\ a' + \alpha' &amp; b' + \beta' &amp; c' + \gamma' \\ a'' + \alpha'' &amp; b'' + \beta'' &amp; c'' + \gamma'' \end{vmatrix} = \begin{vmatrix} a &amp; b &amp; c \\ a' &amp; b' &amp; c' \\ a'' &amp; b'' &amp; c'' \end{vmatrix} + \begin{vmatrix} \alpha &amp; \beta &amp; \gamma \\ \alpha' &amp; \beta' &amp; \gamma' \\ \alpha'' &amp; \beta'' &amp; \gamma'' \end{vmatrix}</math> as a rule for the addition of matrices ; that for their subtraction is of course similar to it.  5. A matrix is not altered but the addition or subtraction of the matrix zero, that is, we have  <math display="block">M \pm O = M.</math> the equation <math>L = M</math>, which expresses that the matrices <math>L, M</math> are equal may also be written in the form <math>L - M = O</math>, i.e. the difference of two matrices is the matrix zero.  6. the equation <math>L = -M</math>, written in the form <math>L + M = O</math>, expresses that the sum of the matrix <math>L, M</math> is equal to the matrix zero, the matrices so related are said to be <i>opposite</i> to each other; in other words, a matrix the terms of which are equal but opposite in sign to the terms of a given matrix is said to be opposite to the given matrix.</p>

Les deux mémoires présentent la définition des éléments neutres 0 et 1, des propriétés de commutativité de l'addition, de l'associativité de l'addition et de la multiplication, et de la distributivité de la multiplication sur l'addition. Eduard Weyr définit la matrice nulle après la définition de l'addition, comme élément neutre ; la même remarque s'applique à la matrice unité [Weyr, 1890, 165], définie comme unique solution de l'équation  $AI = A$  pour tout  $A$ . Cayley, au contraire, définit les matrices nulles et unités avant de donner les règles d'opérations; la définition procède de la représentation sous jacente chez Cayley d'une matrice comme une fonction linéaire : la matrice nulle est la fonction qui transforme les quantités  $(X, Y, Z)$  en  $(0, 0, 0)$  et la matrice unité la fonction qui laissent des quantités inchangées. Cette différence dans l'organisation des propriétés du calcul symbolique manifeste une différence de conception de la notion de matrice : pour Cayley, une matrice est une représentation qui a "la forme d'un carré", c'est une notion qui se "dégage naturellement de la notation abrégée d'un ensemble d'équations linéaires" [Cayley, 1858, 475] ; pour Weyr, c'est un symbole dont il faut définir les propriétés opératoires.

## Encart 13 : la forme canonique de Jordan en 1870.

Le théorème suivant est énoncé par Jordan dans son *Traité des Substitutions* de 1870.

THEOREME. – Soit

$$A = [x, x', \dots, ax+bx'+\dots, a'x+b'x'+\dots, \dots]$$

une substitution linéaire quelconque à coefficients entiers entre  $n$  indices variables chacun de 0 à  $p-1$ ; Soient  $F, F', \dots$  les facteurs irréductibles de la congruence de degré  $n$

$$\begin{vmatrix} a-K & a' & \dots \\ b & b-K & \dots \\ \dots & \dots & \dots \end{vmatrix} \equiv 0 \pmod{p}$$

$l, l', \dots$  leurs degrés respectifs ;  $m, m', \dots$  leurs degrés de multiplicité ;

On pourra remplacer les  $n$  indices indépendants  $x, x', \dots$  par d'autres indices jouissant des propriétés suivantes :

1° Ces indices se partagent en systèmes correspondants aux divers facteurs  $F, F', \dots$  et contenant respectivement  $lm, l'm', \dots$  indices ;

2° Soient  $K_0, K_1, \dots, K_{l-1}$  les racines de la congruence irréductible  $F \equiv 0 \pmod{p}$  ; les  $n$  indices du système correspondant à  $F$  se partagent en  $l$  séries correspondantes aux racines  $K_0, K_1, \dots, K_{l-1}$  ;

3° Les indices de la première série de ce système sont des fonctions linéaires des indices primitifs, dont les coefficients sont des entiers complexes formés avec l'imaginaire  $K_0$  ; ils constituent une ou plusieurs suite  $y_0, z_0, u_0, \dots, y'_0, z'_0, u'_0, \dots$  (\*) telles que  $A$  remplace les indices  $y_0, z_0, u_0, \dots$  d'une même suite respectivement par  $K_0 y_0, K_0(z_0+y_0), K_0(u_0+z_0), \dots$  ;

4° Les indices de la  $r+1^{\text{ième}}$  série sont les fonctions  $y_r, z_r, u_r, \dots; y'_r, z'_r, u'_r, \dots$  respectivement conjuguées des précédentes, que l'on forme en y remplaçant  $K_0$  par  $K_r$  ;  $A$  les remplace respectivement par  $K_r y_r, K_r(z_r+y_r), K_r(u_r+z_r, \dots; \dots)$

Cette forme simple

$$\begin{vmatrix} y_0, z_0, u_0, \dots, y'_0, \dots & K_0 y_0, K_0(z_0+y_0), K_0(u_0+z_0), \dots, K_0 y'_0 \\ y_1, z_1, u_1, \dots, y'_1, \dots & K_1 y_1, K_1(z_1+y_{10}), K_1(u_1+z_1), \dots, K_1 y'_1 \\ \dots & \dots \\ v_0 & K'_0 v_0, \dots \\ \dots & \dots \end{vmatrix}$$

à laquelle on peut ramener la substitution  $A$  par un choix d'indice convenable, sera pour nous sa forme canonique.

[Jordan, 1870, 127].

## Encart 14 : Pincherle et la recherche de sous variétés d'un espace vectoriel invariantes par rapport à une homographie A.

[On] considère les substitutions linéaires comme des opérations (l'auteur dit *opérateurs*) appliquées aux vecteurs de l'espace ordinaire [...]. Au même ordre d'idées, bien que le concept vectoriel n'y soit pas explicitement énoncé, se rattachent d'autres travaux importants : pour n'en citer qu'un, le mémoire de M. Frobenius sur les formes bilinéaires et les nombreuses recherches que ce beau mémoire à inspirées. [...] L'objet de ce mémoire est l'étude des opérations distributives qu'on peut appliquer aux fonctions analytiques, plus particulièrement aux séries de puissances entières et positives de la variable  $x$  [...]. Nous considérons l'ensemble de ces séries comme une variété ou espace (que nous appelons espace fonctionnel) à un nombre infini de dimensions : chaque série est un élément ou point de cet espace, et le système des coefficients de la série peut se regarder comme le système de ses coordonnées. [...] après quelques généralités sur les opérations distributives, on traite des racines de ces opérations, puis des opérations appliquées aux variétés linéaire à un nombre fini de dimensions, et au moyen du concept de racines d'une opération, on retrouve d'une façon tout à fait simple et naturelle la décomposition d'une homographie par les diviseurs élémentaires de Weierstrass.

Etant données l'opération  $A$  et la variété  $V(a_1, a_2, \dots, a_n)$ , invariante par rapport à  $A$  et qui ne renferme pas de racines de  $A$ , existe-t-il, dans cette variété, une fonction  $\omega$  invariante par rapport à  $A$  ?  
[...] Soit donc

$$\omega = h_1 a_1 + h_2 a_2 + \dots + h_n a_n$$

une fonction invariante par rapport à  $A$ ; on aura

$$A(\omega) = z\omega$$

[...] d'où suit pour  $z$  la condition nécessaire et suffisante d'être racine de l'équation caractéristique ou fondamentale de l'opération  $A$  par rapport à la variété donnée

[...] On démontre immédiatement que si  $V_n(a)$  contient une variété  $V_m(b)$ ,

$m < n$ , qui est aussi invariante par rapport à  $A$ , et  $g(t)$  est le premier membre de l'équation fondamentale de  $A$  par rapport à  $V_m(b)$ ,  $g(t)$  est diviseur de  $f(t)$ .

[...] A côté de l'opération  $A$  considérons l'opération

$$(11) B = E_1^{r_1} E_2^{r_2} \dots E_q^{r_q}; (E_i = A - z_i, i = 1, 2, \dots, q).$$

Nous savons que sa variété racine est la somme des variétés racines de  $E_1^{r_1}, E_2^{r_2}, \dots, E_q^{r_q}$ . De plus nous pouvons dire qu'à cette variété appartiennent toutes les fonctions invariantes par rapport à  $A$ .

Nous nous proposons maintenant d'étudier les racines de l'opération  $B$ , et pour cela, celles de l'opération  $E^p$ , où  $E = A - z$ , et  $z$  désigne l'une quelconque des racines de (10).

[...] si l'opération  $E^p = (A - z)^p$  admet, dans la variété  $V_n(a)$ , une racine propre  $\omega^{(p-1)}$  on aura :

$$E(\omega^{(p-1)}) = \omega^{(p-2)}, E(\omega^{(p-2)}) = \omega^{(p-3)}, \dots, E(\omega^{(1)}) = \omega, E(\omega) = 0,$$

d'où l'on tire :

$$(12) \begin{cases} A(\omega) &= z\omega, \\ A(\omega^2) &= z\omega^{(1)} + \omega, \\ \dots &\dots \dots \\ A(\omega^{p-1}) &= z\omega^{(p-1)} + \omega. \end{cases}$$

Les fonctions  $\omega, \omega^{(1)}, \dots, \omega^{(p-1)}$  ainsi déterminées sont linéairement indépendantes [...]. Si donc  $E^p$  admet dans la variété  $V_n(a)$  une racine propre  $\omega^{(p-1)}$ , il existe une variété d'ordre  $p$ ,

$$V_p(\omega, p\omega^{(1)}, \dots, \omega^{(p-1)})$$

également invariante par rapport à  $A$ , contenue dans  $V_n(a)$ , et sur laquelle  $A$  opère la substitution linéaire (12). Le tableau (12) montre que le premier membre de l'équation caractéristique correspondante est

$$g(t) = (t - z)^p,$$

d'où il suit que  $z$  est racine de  $f(t)$  de l'ordre  $p$  au moins.

[...]\*

\*Le lecteur a sans doute déjà remarqué que la théorie des opérations distributives nous a permis d'obtenir le concept des diviseurs élémentaires de Weierstrass, et cela, d'une façon excessivement simple.

[Pincherle, 1899, 341].

## Encart 15 : extraits de différents théorèmes énoncés dans des réseaux de recherches distincts.

# Une méthode de décomposition pour les systèmes d'équations différentielles.

[Sauvage, 1891] :

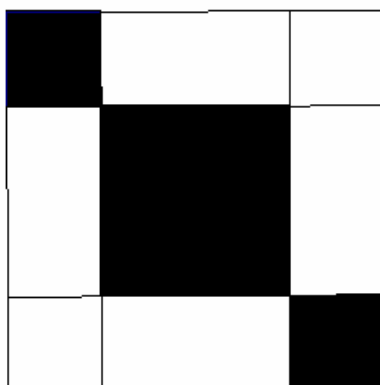


Fig. 2



# Une méthode de décomposition en arithmétique.

[Hensel, 1903, 158-159]

B=	$a_9$	$b_{51}$	$b_{52}$	$c_1$	$a_8$	$b_{11}$	$b_{12}$	$a_5$	$a_1$
	0	$b_{61}$	$b_{62}$	$c_2$	0	$b_{21}$	$b_{22}$	$a_6$	$a_2$
	0	$b_{71}$	$b_{72}$	$c_3$	0	$b_{31}$	$b_{32}$	$a_7$	$a_3$
	0	0	0	$c_4$	0	$b_{41}$	$b_{42}$	0	$a_4$
	0	0	0	0	$a_9$	$b_{51}$	$b_{52}$	$a_8$	$a_5$
	0	0	0	0	0	$b_{61}$	$b_{62}$	0	$a_6$
	0	0	0	0	0	$b_{71}$	$b_{72}$	0	$a_7$
	0	0	0	0	0	0	0	$a_9$	$a_8$
	0	0	0	0	0	0	0	0	$a_9$

### Une méthode de décomposition pour la représentation des systèmes hypercomplexes.

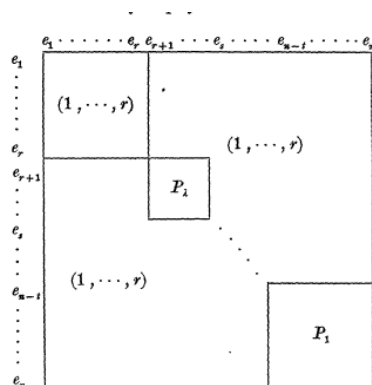
H.E. Hawks [1903, 437-439]

This paper continues and concludes the solution of the general enumeration problem of hypercomplex number-systems which are associative and have a modulus. [...]

Def. 2. a system is reducible if its units may be divided into two or more subsystems [...]

In order to enumerate the various systems of distinct types it is necessary to find a normal form [...].

These theorems we may display in a table as follows



### Une méthode de décomposition pour la représentation des groupes.

[Frobenius, 1903, 307-308].

$$U = \begin{pmatrix} S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & S_k \end{pmatrix}, \quad V = \begin{pmatrix} T_1 & 0 & \dots & 0 \\ 0 & T_2 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & T_k \end{pmatrix}.$$

dann ist

$$S = \begin{pmatrix} U & 0 \\ U_0 & S_0 \end{pmatrix}, \quad T = \begin{pmatrix} V & 0 \\ V_0 & T_0 \end{pmatrix},$$

[...]

Da  $S_x$  die Matrix einer einfachen Gruppe ist [...]